TRANSACTIONS OF

THE ROYAL SOCIETY OF CANADA

SECTION III

CHEMICAL, MATHEMATICAL, AND PHYSICAL SCIENCES

THIRD SERIES—VOLUME XLIX—SECTION III

JUNE, 1955

OTTAWA
THE ROYAL SOCIETY OF CANADA
1955

TRANSACTIONS OF

THE ROYAL SOCIETY
OF CANADA

SECTION III

CHEMICAL, MATHEMATICAL,
AND
PHYSICAL SCIENCES

THIRD SERIES—VOLUME XLIX—SECTION III

JUNE, 1955

OTTAWA
THE ROYAL SOCIETY OF CANADA
1955

CONTENTS

TRANSACTIONS OF THE ROYAL SOCIETY OF CANADA
VOLUME XLIX : SERIES III : JUNE, 1955
SECTION THREE

The Metric of a Conformally Euclidean Space referred to a Subspace[1]

RICHARD BLUM

Presented by R. L. JEFFERY, F.R.S.C.

Vranceanu (**3**; **4**, pp. 343 ff.) has shown that, if in a Euclidean space E_N a subspace V_n with the metric:

$$ds^2 = a_{ij}\, \mathrm{d}x^i\, \mathrm{d}x^j \qquad (i, j, k, \ldots = 1, 2, \ldots, n) \tag{1}$$

is defined by the equations:

$$X^I = f^I(x^1, x^2, \ldots, x^n) \quad (I, J, K, \ldots = 1, 2, \ldots, N) \tag{2}$$

then the metric of E_N can be written

$$\begin{aligned} ds^2 = {} & (a_{ij} - 2x^\alpha b_{\alpha|ij} + x^\alpha x^\beta c_{\alpha\beta|ij})\, dx^i\, dx^j \\ & + 2x^\beta t_{\alpha\beta|i}\, dx^i\, dx^\alpha + \delta_{\alpha\beta}\, dx^\alpha\, dx^\beta. \\ & \qquad (\alpha, \beta, \gamma = n+1, n+2, \ldots, N) \end{aligned} \tag{3}$$

and that this formula is valid in a sufficiently small neighbourhood of V_n. The $b_{\alpha|ij}$ and $c_{\alpha\beta|ij}$ are the fundamental tensors of the second and third kind respectively and $t_{\alpha\beta|i}$ are the "torsions." In the general case (i.e., for a V_n in a V_N) the $b_{\alpha|ij}$ and $t_{\alpha\beta|i}$ are defined by Weingarten's formulas (**1**, pp. 160 f.):

$$f^I_{,ij} = -\left\{ \begin{matrix} I \\ JK \end{matrix} \right\} f^J_{,i} f^K_{,j} + b_{\alpha|ij} \xi^I_\alpha \tag{4}$$

$$\xi^I_{\alpha,i} = -\, b_{\alpha|ij} a^{jk} f^I_{,k} - \left\{ \begin{matrix} I \\ JK \end{matrix} \right\} f^J_{,i}\, \xi^K_\alpha + t_{\beta\alpha|i}\, \xi^I_\beta \tag{5}$$

from which follows by solving for $b_{\alpha|ij}$ and $t_{\alpha\beta|i}$:

$$b_{\alpha|ij} = A_{IJ} f^I_{,ij}\, \xi^J_\alpha + [I, JK]\, \xi^I_\alpha f^J_{,i} f^K_{,j} \tag{4'}$$

$$t_{\alpha\beta|i} = A_{IJ}\, \xi^I_\alpha\, \xi^J_{\beta,i} + [I, JK]\, \xi^I_\alpha\, \xi^J_\beta\, f^K_{,i} \tag{5'}$$

It can easily be verified that the $t_{\alpha\beta|i}$ have the property:

$$t_{\alpha\beta|i} + t_{\beta\alpha|i} = 0 \tag{6}$$

[1]This paper was prepared while the author attended the Summer Research Institute of the Canadian Mathematical Congress.

If V_N is Euclidean (i.e., $A_{IJ} = \delta_{IJ}$) then (4′) and (5′) become:

$$b_{\alpha|ij} = f^I_{.ij}\,\xi^I_\alpha \tag{4″}$$

$$t_{\alpha\beta|i} = \xi^I_\alpha\,\xi^I_{\beta.i} \tag{5″}$$

It is readily recognized that a generalization of formula (3) to a V_N (in which a V_n is imbedded) is not possible, but that it remains feasible if V_N is a conformally euclidean space $(=C_N)$. It turns out that this can be achieved in two ways. If we require that the metric of C_N should (up to a factor) have the same form as (3), then the property (6) of the $t_{\alpha\beta|i}$ has to be abandoned. If, on the other hand, we want to preserve the property (6), the form of the metric of C_N will be slightly changed. In both cases the $b_{\alpha|ij}$ and $t_{\alpha\beta|i}$ will have to be defined by formulas differing from (4′) and (5′).

I. Let C_N be a conformally euclidean space of N dimensions, i.e. a Riemannian space in which a coordinate system $X^1, X^2, \ldots, X^N$, can be found with respect to which the metric of C_N can be written:

$$dS^2 = e^{2\sigma}[(dX^1)^2 + (dX^2)^2 + \ldots + (dX^N)^2]. \tag{I, 1}$$

In this case it is appropriate to call the $X^1, X^2, \ldots, X^N$ a cartesian coordinate system of C_N. The function

$$\sigma = \sigma(X^1, X^2, \ldots, X^N)$$

is assumed to be continuous and with continuous derivatives up to the third order.

Let V_n be a subspace of C_N given by the equations:

$$X^I = f^I(x^1, x^2, \ldots, x^n) \qquad n < N \tag{I, 2}$$

where the functions f^I satisfy the same conditions as σ and the functional matrix $\|f^I_{.i}\|$ is of rank n. It is well known that, in this case, the V_n is a Riemannian space with the positive definite metric:

$$ds^2 = a_{ij}\,dx^i\,dx^j, \tag{I, 3}$$

where

$$a_{ij} = e^{2'\sigma} f^I_{.i} f^I_{.j};\quad {}'\sigma = \sigma(f^1, f^2, \ldots, f^N). \tag{I, 4}$$

Consider now a point $P(x^1, x^2, \ldots, x^n)$ in V_n. At this point we can choose $N - n$ vectors $\vec{\xi}_\alpha$ lying in C_N which are:

(α) Normal to V_n,
(β) Perpendicular to one another,
(γ) Of unit length.

If the components of these vectors are denoted by

$$\xi^I_\alpha = \xi^I_\alpha\,(x^1, x^2, \ldots, x^n)$$

the above conditions can be written:

$$(\mathrm{I}, 5)\qquad \begin{aligned} f^I_{.i}\,\xi^I_\alpha &= 0 \\ \xi^I_\alpha\,\xi^I_\beta &= e^{-2\sigma}\,\delta_{\alpha\beta}\,. \end{aligned}$$

The vectors ξ^I_α define the normal vector space of C_N in P (**2**, p. 320). A point M in this space, considered as a point in C_N, has cartesian coordinates given by the formulas:

$$(\mathrm{I}, 6)\qquad X^I = f^I + x^\alpha \xi^I_\alpha$$

in which it is easily seen that x^α equals the length of the projection of $\vec{PM}$ on $\vec{\xi_\alpha}$.

On the other hand formulas (I, 6) can be considered as a coordinate transformation in C_N, which leads from the $X^1, X^2, \dots, X^N$ to $x^1, \dots, x^n, x^{n+1}, \dots, x^N$. This will be an allowable coordinate transformation in a sufficiently small neighbourhood of V_n since its Jacobian is different from zero on V_n. (It can be easily verified that the square of the Jacobian equals on V_n, up to a non-vanishing factor, the determinant $|a_{ij}|$.)

From (I, 6) we obtain by differentiating:

$$(\mathrm{I}, 7)\qquad dX^I = (f^I_{.i} + x^\alpha\,\xi^I_{\alpha,i})\,dx^i + \xi^I_\alpha\,dx^\alpha$$

and by substituting in (I, 1):

$$(\mathrm{I}, 8)\qquad \begin{aligned} dS^2 = e^{2\sigma}[(f^I_{.i}f^I_{.j} + 2x^\alpha f^I_{.i}\,\xi^I_{\alpha,j} + x^\alpha x^\beta\,\xi^I_{\alpha,i}\,\xi^I_{\beta,j})\,dx^i dx^j \\ + 2x^\beta\,\xi^I_\alpha\,\xi^I_{\beta,i} dx^i dx^\alpha + e^{-2\sigma}\,\delta_{\alpha\beta}\,dx^\alpha\,dx^\beta] \end{aligned}$$

where, of course, in σ the X^I are given by formulas (I, 6).

From (I, 4) we have:

$$(\mathrm{I}, 4')\qquad f^I_{.i}f^I_{.j} = e^{-2\sigma}\,a_{ij}\,.$$

By differentiating (I, 5) we obtain:

$$(\mathrm{I}, 9)\qquad \begin{aligned} f^I_{.ij}\,\xi^I_\alpha + f^I_{.i}\,\xi^I_{\alpha,j} &= 0 \\ \xi^I_\alpha\,\xi^I_{\beta,i} + \xi^I_\beta\,\xi^I_{\alpha,i} &= -\,2e^{-2\sigma}\,\delta_{\alpha\beta}\,'\sigma_{,i}\,. \end{aligned}$$

We define now three sets of tensors in V_n by the formulas:

$$(\mathrm{I}, 10)\qquad \begin{aligned} \bar{b}_{\alpha|ij} &= e^{2\sigma} f^I_{.ij}\,\xi^I_\alpha = -\,e^{2\sigma} f^I_{.i}\,\xi^I_{\alpha,j} \\ \bar{c}_{\alpha\beta|ij} &= e^{2\sigma}\,\xi^I_{\alpha,i}\,\xi^I_{\beta,j} \\ \bar{l}_{\alpha\beta|i} &= e^{2\sigma}\,\xi^I_\alpha\,\xi^I_{\beta,i}\,. \end{aligned}$$

It is obvious that

$$(\mathrm{I}, 11)\qquad \begin{aligned} \bar{b}_{\alpha|ij} &= \bar{b}_{\alpha|ji}, \\ \bar{c}_{\alpha\beta|ij} &= \bar{c}_{\alpha\beta|ji} = \bar{c}_{\beta\alpha|ij}\,. \end{aligned}$$

Furthermore, from (I, 8) we find that:

$$(\mathrm{I}, 12) \qquad \bar{t}_{\alpha\beta|i} + \bar{t}_{\beta\alpha|i} = -2\delta_{\alpha\beta}{}'\sigma_{,i}.$$

Because of (I, 4′), (I, 10) and (I, 9), formula (I, 8) can now be written:

$$(\mathrm{I}, 13) \quad dS^2 = e^{2\sigma - 2'\sigma}[(a_{ij} - 2x^\alpha \bar{b}_{\alpha|ij} + x^\alpha x^\beta \bar{c}_{\alpha\beta|ij})\, dx^i dx^j + 2x^\beta \bar{t}_{\alpha\beta|i} dx^i dx^\alpha + \delta_{\alpha\beta} dx^\alpha dx^\beta].$$

This is the desired form of the metric of C_N. The $\bar{t}_{\alpha\beta|i}$ are no longer skew-symmetric, but satisfy the more general relation (I, 12).

If we desire to preserve the skew-symmetry of the $t_{\alpha\beta|i}$ the condition (γ) has to be replaced by the condition:

(γ') Of equal length $e^{'\sigma}$.

This will lead through obvious stages, based on an identical line of thought as above, to the following metric for C_N:

$$(\mathrm{I}, 13_I) \quad dS^2 = e^{2\sigma - 2'\sigma}[(a_{ij} - 2x^\alpha \bar{\bar{b}}_{\alpha|ij} + x^\alpha x^\beta \bar{\bar{c}}_{\alpha\beta|ij})\, dx^i dx^j + 2x^\beta \bar{\bar{t}}_{\alpha\beta|i} dx^i dx^\alpha + e^{2'\sigma} \delta_{\alpha\beta} dx^\alpha dx^\beta].$$

In this formula the $\bar{\bar{b}}_{\alpha|ij}$, $\bar{\bar{c}}_{\alpha\beta|ij}$ and $\bar{\bar{t}}_{\alpha\beta|i}$ are defined in terms of a set of vectors η^I_α in exactly the same way as the $\bar{b}_{\alpha|ij}$, $\bar{c}_{\alpha\beta|ij}$ and $\bar{t}_{\alpha\beta|i}$ in terms of the ξ^I_α through formulas (I, 10), and where, in addition:

$$\eta^I_\alpha = e^{'\sigma}\, \xi^I_\alpha.$$

That the $\bar{b}_{\alpha|ij}$ and $\bar{t}_{\alpha\beta|i}$, thus defined, do not coincide with the $b_{\alpha|ij}$ and $t_{\alpha\beta|i}$ as given by (4′) and (5′) can easily be recognized by substituting in the formulas:

$$A_{IJ} = e^{2\sigma}\, \delta_{IJ}$$

We then obtain:

$$(\mathrm{I}, 14) \qquad \begin{aligned} b_{\alpha|ij} &= \bar{b}_{\alpha|ij} - {}'\sigma_\alpha\, a_{ij}; \quad \sigma_\alpha = \sigma_{,I}\xi^I_\alpha \\ t_{\alpha\beta|i} &= \bar{t}_{\alpha\beta|i} + \delta_{\alpha\beta}{}'\sigma_{,i}. \end{aligned}$$

On the other hand the following relations can be established:

$$(\mathrm{I}, 15) \qquad \begin{aligned} b_{\alpha|ij} &= e^{'\sigma}\, \bar{b}_{\alpha|ij} \\ t_{\alpha\beta|i} &= e^{2'\sigma}\, (\bar{t}_{\alpha\beta|i} + \delta_{\alpha\beta}{}'\sigma_{,i}) \end{aligned}$$

from which we have:

$$(\mathrm{I}, 16) \qquad \begin{aligned} \bar{\bar{b}}_{\alpha|ij} &= e^{'\sigma}(b_{\alpha|ij} + {}'\sigma_\alpha\, a_{ij}) \\ \bar{\bar{t}}_{\alpha\beta|i} &= e^{2'\sigma} t_{\alpha\beta|i}. \end{aligned}$$

It follows therefore that in general none of the three sets of $b_{\alpha|ij}$ and $t_{\alpha\beta|i}$ coincide.

The importance of formula (I, 13) lies in the fact that the metric

$$
\begin{aligned}
d\bar{S}^2 = {} & (a_{ij} - 2x^\alpha \bar{b}_{\alpha|ij} + x^\alpha x^\beta \bar{c}_{\alpha\beta|ij})\, dx^i dx^j \\
& + 2x^\beta \bar{t}_{\alpha\beta|i} dx^i dx^\alpha + \delta_{\alpha\beta} dx^\alpha dx^\beta
\end{aligned}
\tag{I, 17}
$$

is also conformally-euclidean and that therefore the components of its conformal curvature tensor C^I_{JKL} (**1**, p. 90) are zero. The equations:

$$
{}'C^I_{JKL} = 0 \tag{I, 18}
$$

turn out to be the equations of Gauss, Codazzi, and Kuehne of the V_n in C_N, thus leading to an alternate proof of the fundamental theorem of subspaces of a conformally euclidean space, which states that the *necessary and sufficient* condition for the a_{ij}, $\bar{b}_{\alpha|ij}$ and $\bar{t}_{\alpha\beta|i}$ to be respectively the fundamental tensors of the first and second kind and the torsions of a V_n imbedded in C_N, is that they should satisfy the equations (I, 18) (**5**, pp. 437–449) (**6**, pp. 338–342).

An investigation of these equations and their consequences will be the subject of a future paper.

References

1. L. P. Eisenhart, *Riemannian geometry* (Princeton, 1926).

2. A. Fialkow, *Conformal differential geometry of a subspace*, Trans. Am. Math. Soc., *56* (1944), 309–433.

3. G. Vranceanu, *Sulle condizioni di rigidita di una V_m in un S_n*, Rend. Lincei, *11* (1930), 375–389.

4. ———, *Leçons de géométrie différentielle*, vol. I (Bucharest, 1947).

5. K. Yano and Y. Muto, *Sur le théorème fondamental dans la géométrie conforme des sous-espaces riemanniens*, Proc. Phys. Math. Soc. Japan, *24* (1942), 437–449.

6. ———, *Note sur le théorème fondamental dans la géometrie conforme des sous-espaces riemanniens*, Proc. Japan Acad., *12* (1946), 338–342.

TRANSACTIONS OF THE ROYAL SOCIETY OF CANADA
VOLUME XLIX : SERIES III : JUNE, 1955
SECTION THREE

A Quasi-Linear Boundary Value Problem[1]

G. F. D. DUFF

Presented by R. L. JEFFERY, F.R.S.C.

The existence theorem established in this note relates to a second-order quasi-linear elliptic partial differential equation together with a quasi-linear boundary condition. The proof, which is an extension of the Bergman-Schiffer method of integro-differential equations, suggests a construction of the solution by successive approximations, each of which involves the solution of a linear problem. To avoid complicated estimates of the singularities of certain domain functionals, I have assumed that the differential equation and the domain are analytic, in a sense to be made precise later.

Non-linear equations with linear boundary conditions have been studied by many writers, most attention being paid to the Dirichlet problem wherein values of the solution on the boundary are assigned **(1, 6)**. A non-linear boundary condition for the Laplace equation was studied by Carleman **(4)** and later by Lichtenstein **(6)**, using a method of fundamental solutions which is restricted to linear differential equations. In the present problem it is necessary to study linear boundary conditions of the Neumann and Robin type, and the theorem includes these as special cases. As in most non-linear problems, some restriction must be made to secure an *a priori* upper bound for the solution.

1. Formulation of the problem. Let D be a compact domain of a space of N dimensions with coordinate variables x^i $(i = 1, \ldots, N)$, and let the boundary B of D be an analytic hypersurface in the space. Consider the equation

$$\Delta u = F(P, u), \tag{1.1}$$

where

$$\Delta u = \frac{1}{\sqrt{a}} \frac{\partial}{\partial x^i}\left(\sqrt{a}\, a^{ik} \frac{\partial u}{\partial x^k}\right) \tag{1.2}$$

is the Laplacian based on a non-singular analytic and symmetric contravariant tensor field a^{ik} defined in $D + B$.

[1]This paper was prepared at the 1955 Summer Research Institute of the Canadian Mathematical Congress. The author's thanks are due the National Research Council of Canada for a fellowship held there.

The (non-linear) function $F(P, u)$ shall be an analytic function of co-ordinates, represented by P, and of the unknown function u, satisfying the conditions

$$F(P, 0) = 0, \tag{1.3}$$

$$F_u(P, u) \equiv \frac{\partial F}{\partial u}(P, u) \geqslant \delta > 0. \tag{1.4}$$

Here δ is an arbitrary fixed positive number. Condition (1.3), which can be modified without noticeably changing the narure of the proof, is imposed to gain a starting point for the construction by successive approximations. Condition (1.4), on the other hand, is needed to secure an *a priori* upper bound for the solution and cannot be relaxed. The requirement that F_u should be bounded away from zero is stronger than the non-negative property used by Bergman and Schiffer for the Dirichlet problem **(1)**.

The quasi-linear boundary condition has the form

$$\frac{\partial u}{\partial n} + B(p, u) = f(p). \tag{1.5}$$

We introduce here the convention of denoting points of B by small Latin letters. The normal derivative in (1.5) is that appropriate to the Riemannian geometry of the associate tensor a_{ik} of the above coefficient tensor a^{ik}. The function $B(p, u)$ shall satisfy the conditions

$$B(p, 0) \geqslant 0, \tag{1.6}$$

$$B_u(p, u) > 0, \tag{1.7}$$

which are analogous to (1.3) and (1.4), respectively. The datum function $f(p)$ shall be once continuously differentiable.

The main result can now be stated.

THEOREM. *There exists a unique solution of* (1.1) *which satisfies the boundary condition* (1.5), *and this solution is analytic in D.*

The uniqueness proof, which is short, will be given first, and then the integro-differential equation method will be developed. We remark that the analyticity of the solution follows from standard theorems under these hypotheses.

2. Proof of uniqueness. Let $u_1(P)$ and $u_2(P)$ be solutions of the problem, and set

$$w(P) = u_2(P) - u_1(P).$$

If $u_2(P) \not\equiv u_1(P)$, there exist regions D_1 and D_2 not both null, and such that

$$w(P) > 0 \qquad \text{in } D_1 \tag{2.1}$$

and

$$w(P) < 0 \qquad \text{in } D_2. \tag{2.2}$$

Suppose D_1 is not null. Then, from (1.5), we have

$$\frac{\partial w}{\partial n} = B(p, u_1) - B(p, u_2) \leqslant 0 \tag{2.3}$$

on the portion B_1 of B adjoining D_1, and

$$w = 0$$

on the remaining part S of the boundary of D_1.

We now apply Green's first formula to the function w on D_1, and find

$$\int_{D_1} w\, \Delta w\, dV + \int_{D_1} (\Delta w)^2 dV = \int_{B_1 + S} w \frac{\partial w}{\partial n}\, dS. \tag{2.4}$$

Since

$$\Delta w = F(P, u_2) - F(P, u_1) \geqslant 0$$

in view of (1.4), and w is non-negative in D_1, the left side of (2.4) is non-negative. However the right side is non-positive in consequence of (2.1) and (2.3). Each side of (2.4) must therefore vanish. This requires that

$$(\nabla w)^2 \equiv a^{ik} \frac{\partial w}{\partial x^i} \frac{\partial w}{\partial x^k} = 0,$$

from which it follows that w = constant in D_1. This constant can only be zero since $w = 0$ on S. Thus D_1 must be null. Similarly D_2 must be null. Hence $w \equiv 0$ and $u_2 \equiv u_1$, which proves the uniqueness of the solution.

3. The equations of variation. If the datum function is zero, the problem has the solution $u \equiv 0$. We therefore consider the problem in which the data are $tf(p)$ for $0 \leqslant t \leqslant 1$, and calculate the growth of the solution as t increases to the value unity. Thus, let $u(P, t)$ be the (unique) solution of the equation

$$\Delta u(P, t) = F(P, u(P, t)) \tag{3.1}$$

with

$$\frac{\partial u(p, t)}{\partial n} + B(p, u(p, t)) = tf(p). \tag{3.2}$$

Now let us define

$$v(P, t) = \frac{\partial u(P, t)}{\partial t} \tag{3.3}$$

and find the equation and boundary condition satisfied by $v(P, t)$. Differentiating (3.1) with respect to t, we have the linear equation of variation for v:

$$\Delta v(P, t) = F_u(P, u(P, t))\, v(P, t). \tag{3.4}$$

From (3.2) we derive the boundary condition

$$\frac{\partial v(p,t)}{\partial n} + B_u(p, u(p,t))\, v(p,t) = f(p), \tag{3.5}$$

which is a linear condition of the Robin type, if $u(P, t)$ is assumed known. The theory of the linear equation shows **(2, 3)** that

$$v(P,t) = \int_B R(P,Q,t) f(q)\, dS_q, \tag{3.6}$$

where $R(P, Q, t)$ is the Robin function for the equation (3.4) in which $F_u(P, u)$ is regarded as known. The condition (1.4) ensures that this solution exists and is unique. Thus the Robin function can be defined as the fundamental solution of (3.4) which satisfies the homogeneous boundary condition corresponding to (3.5).

As t varies, $u(P, t)$ changes and this affects the Robin function in two ways: through the variation of the coefficient F_u in (3.4) and the function B in (3.5). The resulting variation of the Robin function may be calculated with the help of Green's formula. Since the work is straightforward and follows the method of **(2)**, it is omitted, and the final result only will be stated. For the coefficient F_u the change of $R(P, Q, t)$ is

$$\delta_1 R(P,Q,t) = -\int_D F_{uu}(Z,u) \frac{\partial u(Z,t)}{\partial t} R(P,Z,t)\, R(Q,Z,t)\, dV\, \delta t, \tag{3.7}$$

while the variation arising from the change of B_u in (3.5) is

$$\delta_2 R(P,Q,t) = -\int_B B_{uu}(z,u) \frac{\partial u(z,t)}{\partial t} R(P,z,t)\, R(Q,z,t)\, dS_z\, \delta t. \tag{3.8}$$

From (3.3), (3.6), (3.7), and (3.8) we find for $u(P, t)$ and $R(P, Q, t)$ the following system of integro-differential equations:

$$\frac{\partial u(P,t)}{\partial t} = \int R(P,q,t) f(q)\, dS_q, \tag{3.9}$$

$$\begin{aligned} \frac{\partial R(P,Q,t)}{\partial t} = &-\int_D F_{uu}(Z,u) \frac{\partial u(Z,t)}{\partial t} R(P,Z,t)\, R(Q,Z,t)\, dV_Z \\ &-\int_B B_{uu}(z,u) \frac{\partial u(z,t)}{\partial t} R(P,z,t)\, R(Q,z,t)\, dS_z. \end{aligned} \tag{3.10}$$

The initial conditions for this system are easily seen to be

$$u(P,0) \equiv 0, \tag{3.11}$$

while $R(P, Q, 0)$ is the Robin function for the equation $\Delta w = 0$ with the boundary condition

$$\frac{\partial w}{\partial n} + B_u(p,0)\, w = 0. \tag{3.12}$$

This harmonic Robin function exists and can be explicitly calculated in simple cases.

That the solution $u(P, t)$ of (3.9), (3.10) is unique will appear from the proof of its existence. Since $u(P, t)$ satisfies for each t the system (3.1) and (3.2), it is seen that $u(P, 1)$ will be a solution of the original problem.

Bergman and Schiffer **(1)** remark that if the solutions of (1.1) are regarded as points of a non-linear functional manifold, and solutions v of the equation of variation as points of the tangent space at a point u of the manifold, then as t varies, a curve is traced on the non-linear manifold, with the "direction" in the tangent space at each point being given by (3.9). The equation (3.10) describes the change of the tangent space as t varies. The two formulae together are analogous to the Serret-Frenet formulae for a space curve, as treated in differential geometry.

4. Properties of the domain functionals. The system (3.9), (3.10) will be solved by successive approximation. Now to show that this process converges, certain *a priori* limitations on $u(P, t)$ and $R(P, Q, t)$ are needed. We now develop these.

The Robin function $R(P, Q, t)$ is non-negative **(2, 3)**. Thus from (3.7) we see that if the coefficient F_u of (3.4) increases, then $R(P, Q, t)$ decreases. Similarly if B_u in (3.5) increases, $R(P, Q, t)$ decreases. Thus $R(P, Q, t)$ is a monotone decreasing functional of F_u and B_u, both quantities being non-negative. Now let $N_\delta(P, Q)$ denote the Neumann function of the domain D for the equation $\Delta w = \delta w$, and let $\delta > 0$ be the lower bound of F_u as in (1.4). We clearly have

LEMMA 4.1 $$R(P, Q, t) \leqslant N_\delta(P, Q).$$

This is a uniform bound independent of t. We note that δ must be chosen positive since the harmonic Neumann function has a special construction and the above inequalities no longer apply.

To establish a similar bound for $v(P, t)$, we note that according to (3.6), $v(P, t)$ is dominated by the integral

$$\int_B N_\delta(P, q)\, |f(q)|\, dS_q.$$

Let M be the maximum value of the datum function $|f(p)|$. We shall prove

LEMMA 4.2. *If $q \geqslant \delta > 0$, then the solution w of $\Delta w = qw$ satisfying $\partial w/\partial n = f$, $|f| \leqslant M$, on B, satisfies the estimate*

$$|w(P)| \leqslant KM, \qquad P \text{ in } D, \tag{4.1}$$

where K is a constant independent of M and depending only on δ: $K = K(\delta)$.

The solution $w(P)$ is given by

$$w(P) = \int_B R_q(P, r)\, f(r)\, dS_r,$$

and

$$|w(P)| \leqslant \int_B N_\delta(P, r)\, |f(r)|\, dS_r$$

$$(4.2) \qquad \leqslant M \int_B N_\delta(P, r)\, dS_r$$

$$= M \int_B K_\delta(P, r)\, dS_r,$$

where

$$(4.3) \qquad K_\delta(P, R) = N_\delta(P, R) - G_\delta(P, R)$$

is the Bergman kernel function for $\Delta w = \delta w$ **(2, 3)**. This kernel is regular except when P and Q tend to the same point on the boundary, and so $K(P, Q)$ is uniformly bounded for P or Q in $D - N_\epsilon(B)$. Here $N_\epsilon(B)$ denotes an ϵ-neighbourhood of the boundary.

Since the equation $\Delta w = \delta w$ and the boundary are analytic, we may set up a system of coordinates one of which measures inward distance from B. Applying the Cauchy-Kowalewski theorem **(5**, p. 283**)** we may conclude that there exists in some $N_{2\epsilon}(B)$ an analytic solution w_0 of the equation

$$(4.4) \qquad \Delta w_0 - \delta w_0 = 0$$

which satisfies the boundary conditions

$$(4.5) \qquad w_0(p) = 0, \quad \frac{\partial w_0(p)}{\partial n} = 1.$$

We extend the definition of w_0 to the whole domain D in such a way that $w_0(P)$ is C^∞ in D and satisfies (4.4) and (4.5) in $N_\epsilon(B)$. We now apply Green's formula to the kernel function and w_0. Thus

$$\int_B \left[K_\delta(P, q) \frac{\partial w_0(q)}{\partial n} - w_0(q) \frac{\partial K_\delta(P, q)}{\partial n} \right] dS_q$$

$$= \int_D [(\Delta - \delta)\, w_0(Q)\, K_\delta(P, Q) - (\Delta - \delta)\, K_\delta(P, Q) \,.\, w_0(Q)]\, dV_Q.$$

Since $\Delta K = \delta K$, and $\Delta w_0 = \delta w_0$ in $N_\epsilon(B)$, and in view of (4.5), we find

$$(4.6) \qquad \int_B K_\delta(P, q)\, dS_q = \int_{D - N_\epsilon(B)} K_\delta(P, Q)(\Delta - \delta)\, w_0(Q)\, dV_Q.$$

The function $(\Delta - \delta)\, w_0(Q)$ is C^∞ and hence bounded in $D + B$ and $K_\delta(P, Q)$ is uniformly bounded for $P \in D$ and $Q \in D - N_\epsilon(B)$. Thus the right side of (4.6) is uniformly bounded by a constant K independent of P. Combining this with (4.2) we see that the lemma is proved.

For $0 \leqslant t \leqslant 1$, the function $v(P, t) = u_t(P, t)$ satisfies the conditions of Lemma 4.2, and we conclude first that $|v(P, t)| \leqslant KM$ and second that $|u(P, t)| \leqslant KMt$. This furnishes the required uniform bound

$$(4.7) \qquad |u(P, t)| \leqslant KM, \qquad 0 \leqslant t \leqslant 1.$$

5. Solution by successive approximation. The starting point $u_0(P, t)$ is that given by (3.11) and (3.12). For subsequent stages, we define $u_n(P, t)$ as follows:

$$u_n(P, 0) = 0, \tag{5.1}$$
$$\frac{\partial u_n(P, t)}{\partial t} = \int_B R_{n-1}(P, q, t) f(q) \, dS_q.$$

Here $R_n(P, Q, t)$ is the Robin function for the differential equation

$$\Delta w = F_u(P, u_n(P, t)) \, w, \tag{5.2}$$

with the boundary condition

$$\frac{\partial w}{\partial n} + B_u(P, u_n(P, t)) \, w = 0. \tag{5.3}$$

Starting with $u_0(P, t) \equiv 0$, we calculate $R_0(P, Q, t)$, $u_1(P, t)$, $R_1(P, Q, t)$, and so on, in succession.

From Green's formula we can find the following "finite difference" analogue of (3.7) and (3.8):

$$\begin{aligned} R_n(P, Q, t) &- R_{n-1}(P, Q, t) \\ = &- \int_D [F_u(Z, u_n) - F_u(Z, u_{n-1})] \, R_n(P, Z, t) \, R_{n-1}(Z, Q, t) \, dV_Z \\ &- \int_B [B_u(z, u_n) - B_u(z, u_{n-1})] \, R_n(P, z, t) \, R_{n-1}(z, Q, t) \, dS_z. \end{aligned} \tag{5.4}$$

We now define

$$d_n(t) = \max_{P \epsilon D+B} |u_n(P, t) - u_{n-1}(P, t)|. \tag{5.5}$$

Because of the *a priori* bound (4.7) we can state that $F_{uu}(P, u)$ and $B_{uu}(P, u)$ are bounded. Thus, let

$$|F_{uu}(P, u)| \leqslant F, \quad |B_{uu}(P, u)| \leqslant B,$$

when (4.7) holds. Then

$$|F_u(Z, u_n) - F_u(Z, u_{n-1})| \leqslant F \, d_n(t),$$

and

$$|B_u(z, u_n) - B_u(z, u_{n-1})| \leqslant B \, d_n(t).$$

From (5.4) and Lemma 4.1 we now obtain

$$|R_n(P, Q, t) - R_{n-1}(P, Q, t)| \leqslant d_n(t) \, \mathcal{N}(P, Q), \tag{5.6}$$

where

$$\mathcal{N}(P, Q) = F \int_D N_\delta(P, Z) \, N_\delta(Q, Z) \, dV_Z + B \int_B N_\delta(P, z) \, N_\delta(Q, z) \, dS_z. \tag{5.7}$$

We now have

$$\left|\frac{\partial u_{n+1}(P,t)}{\partial t} - \frac{\partial u_n(P,t)}{\partial t}\right|$$

$$(5.8) \qquad < \int_B |R_n(P,z,t) - R_{n-1}(P,z,t)|\, |f(z)|\, dS_z$$

$$\leqslant M\, d_n(t) \int_B \mathcal{N}(P,z)\, dS_z,$$

and so must find a bound for the integral of $\mathcal{N}(P, z)$ appearing on the right.

LEMMA 5.1 $$\int_B \mathcal{N}(P,z)\, dS_z < A,$$

where A is a constant independent of P.

To prove this we use the technique of Lemma 4.2, with some further elaborations. We note that

$$(5.9) \qquad (\Delta_Q - \delta)\, \mathcal{N}(P,Q) = -FN_\delta(P,Q),$$

and

$$(5.10) \qquad \frac{\partial \mathcal{N}(P,q)}{\partial n_q} = B\, N_\delta(P,q).$$

These properties follow immediately from standard theorems concerning the Neumann function. **(3)**. Let $w_1(P)$ be the function defined by the conditions $\Delta w_1 = \delta w_1$ and $\partial w_1/\partial n = 1$. Applying Green's theorem, we have

$$(5.11) \qquad \int_B \left[\mathcal{N}(P,z)\frac{\partial w_1(z)}{\partial n} - w_1(z)\frac{\partial \mathcal{N}(P,z)}{\partial n}\right] dS_z$$

$$= \int_D [(\Delta - \delta)\, w_1(Q) \,.\, \mathcal{N}(P,Q) - (\Delta - \delta)\mathcal{N}(P,Q) \,.\, w_1(Q)]\, dV_Q.$$

Taking account of the definition of w_1, and of (5.9) and (5.10), we see that this simplifies to

$$\int_B \mathcal{N}(P,z)\, dS_z$$

$$(5.12) \qquad = B\int_B w_1(q)\, N(P,q)\, dS_q + F\int_D N(P,Q)\, w_1(Q)\, dV_Q$$

$$= B\, I_1 + F\, I_2,$$

say. To estimate I_1, we note that $|w_1| \leqslant K(\delta)$, by Lemma 4.2, and that the integral I_1 is a solution of $\Delta I_1 = \delta I_1$, with normal derivative having the values of w_1 on B. By a second application of Lemma 4.1, we thus find

$$(5.13) \qquad |I_1| \leqslant K(\delta)^2.$$

For the estimation of I_2 we use the Neumann function $N_{\frac{1}{2}\delta}(P, Q)$ of $\Delta u = \frac{1}{2}\delta u$; this function satisfies the equation

$$(5.14) \qquad (\Delta - \delta)\, N_{\frac{1}{2}\delta}(P, Q) = -\tfrac{1}{2}\delta N_{\frac{1}{2}\delta}(P, Q),$$

and moreover it dominates $N_\delta(P, Q)$ since the Neumann function decreases when δ increases. We note that

$$w_1(P) = \int_B N_\delta(P, q)\, dS_q > 0$$

since the Neumann function is non-negative **(2)**. Thus I_2 is also non-negative and we have

$$\begin{aligned} I_2 &= \int_D N_\delta(P, Q)\, w_1(Q)\, dV_Q \\ &< \int_D N_{\frac{1}{2}\delta}(P, Q)\, w_1(Q)\, dV_Q \\ &= -\frac{2}{\delta}\int_D (\Delta - \delta)\, N_{\frac{1}{2}\delta}(P, Q)\,.\, w_1(Q)\, dV_Q, \end{aligned}$$

by (5.14). If now Green's formula is applied, we find that the last integral above is equal to

$$-\frac{2}{\delta}\int_D N_{\frac{1}{2}\delta}(P, Q)(\Delta - \delta)\, w_1(Q)\, dV_Q - \frac{2}{\delta}\int_B \left[\frac{\partial N_{\frac{1}{2}\delta}}{\partial n} w_1 - N_{\frac{1}{2}\delta}\frac{\partial w_1}{\partial n}\right] dS_q,$$

and the volume integral vanishes according to the definition of $w_1(Q)$. Likewise the normal derivative of $N_{\frac{1}{2}\delta}(P, Q)$ is zero. Thus finally

$$(5.15) \qquad I_2 < \frac{2}{\delta}\int_B N_{\frac{1}{2}\delta}(P, q)\, dS_q$$

since $\partial w_1/\partial n = 1$.

The proof of Lemma 4.1 now shows that the integral in (5.15), which is the solution of

$$\Delta w = \tfrac{1}{2}\delta w$$

having normal derivative unity, is bounded in absolute value by the constant $K(\frac{1}{2}\delta)$. Hence

$$(5.16) \qquad |I_2| < \frac{2}{\delta} K(\tfrac{1}{2}\delta).$$

To complete the proof of Lemma 5.1 we need only choose

$$A = BK(\delta)^2 + 2FK(\tfrac{1}{2}\delta)/\delta.$$

Returning to equation (5.8), we now find

$$(5.17) \qquad \left|\frac{\partial u_{n+1}(P, t)}{\partial t} - \frac{\partial u_n(P, t)}{\partial t}\right| < MAd_n(t).$$

Integrating from 0 to t, and noting that $u_n(P, 0) = 0$, we see that

$$|u_{n+1}(P, t) - u_n(P, t)| \leqslant MA \int_0^t d_n(\tau)\, d\tau. \tag{5.18}$$

Since the quantity on the right in (5.18) is independent of P, we see by (5.5) that

$$d_{n+1}(t) \leqslant MA \int_0^t d_n(\tau)\, d\tau. \tag{5.19}$$

Since $u_0(P, 0) \equiv 0$, we have

$$\begin{aligned} d_1(t) &= \max_P |u_1(P, t)| \\ &\leqslant \max_P \left[\int_B R_0(P, q, t) |f(q)|\, dS_q \right] \\ &\leqslant KM, \end{aligned} \tag{5.20}$$

by Lemmas 4.1 and 4.2. It follows quickly that $d_2(t) \leqslant KM^2At$, and by induction, that

$$d_n(t) \leqslant \frac{K\, M^n A^{n-1} t^{n-1}}{(n-1)!}, \tag{5.21}$$

in the usual manner. The right side of (5.21) is independent of P, and is a term of an absolutely convergent series. Consequently the series $\sum_n d_n(t)$ converges for all t, and so, therefore, does the series

$$\sum_n [u_{n+1}(P, t) - u_n(P, t)]$$

converge uniformly in P and absolutely. Let $u(P, t)$ denote the sum of the series.

From (5.6) we see that the sequence of Robin functions $R_n(P, Q, t)$ converges to a limit $R(P, Q, t)$. Since $N(P, Q)$ may tend to infinity as P tends to Q, the convergence is uniform, for fixed Q, in any compact region for P which does not contain Q. Thus the limit $R(P, Q, t)$ is continuous for $P \neq Q$.

From (3.7) and (3.8) we see that for $P \neq Q$, the Robin function depends continuously on the coefficients of the differential equation and boundary condition. Thus the limit of the Robin functions $R_n(P, Q, t)$ is the Robin function of the limit of equation (5.2) and boundary condition (5.3). That is, $R(P, Q, t)$ is the Robin function for

$$\Delta v = F_u(P, u(P, t))\, v,$$

with boundary condition

$$\frac{\partial v}{\partial n} + B_u(p, u(p, t))\, v = 0.$$

From (3.7) and (3.8) it follows that the rate of change of $R(P, Q, t)$ with t is given by the right side of (3.10).

From (5.17) we observe that the derivatives $\partial u_n(P, t)/\partial t$ converge uniformly and absolutely. In the limit as $n \to \infty$, (5.1) takes the form (3.9). Thus both of the basic integro-differential equations are satisfied by $u(P, t)$ and $R(P, Q, t)$. For $t = 0$ we see from (5.1) that (3.11) and (3.12) are satisfied. Therefore $u(P, 1)$ is a solution of the original problem. We have already proved that such a solution is unique; it follows from standard theory that $R(P, Q, 1)$ is also unique.

6. Concluding remarks. This theorem includes as special cases the linear Neumann and Robin boundary conditions usually treated in potential theory. However the type of boundary condition can be generalized still more. A general non-linear boundary condition has the form

$$B\left(p, \frac{\partial u}{\partial n}, u, f(p)\right) = 0.$$

The restrictions needed in order to apply the preceding theory are

$$B_{u_n} \neq 0; \quad B_u B_{u_n} \geqslant 0; \quad |B_f| < C|B_{u_n}|.$$

The only part of the proof in which the analytic nature of the boundary or the equation has been used is in Lemma 4.1. If the integral on the right of (4.6) can be shown bounded with respect to P in any other way then the assumption of analyticity can be omitted. It seems likely that if $w_0(P)$ is so defined that $(\Delta - \delta)\, w_0(P)$ tends to zero sufficiently rapidly as $P \to B$, this result would hold.

References

1. S. Bergman and M. Schiffer, *Existence theorems for some quasi-linear partial differential equations*, ONR report (1952).
2. ———, *Kernel functions in mathematical physics* (New York, 1953).
3. ———, *Kernel functions in the theory of partial differential equations of elliptic type*, Duke Math. J., *15* (1948), 535–566.
4. T. Carleman, *Ueber eine Nichtlineare Randwertaufgabe*, Math. Zeit., *9* (1921), 35–43.
5. E. Goursat, *Course of analysis*, vol. II, part II (Boston, 1917).
6. L. Lichtenstein, *Nichtlineare Integralgleichungen* (Berlin, 1931).

TRANSACTIONS OF THE ROYAL SOCIETY OF CANADA
VOLUME XLIX : SERIES III : JUNE, 1955
SECTION THREE

On Commuting Automorphisms of Rings[1]

NATHAN DIVINSKY

Presented by R. L. JEFFERY, F.R.S.C.

In the theory of crossed extension algebras there arises an automorphism T of an associative algebra such that

$$x \cdot x^T = x^T \cdot x$$

for every x in the algebra. We shall call such an automorphism a commuting automorphism. If in addition, $T \neq I$, the identity automorphism, we shall say that T is a non-trivial commuting automorphism.

We shall prove as our main result of this note that if a simple associative ring A with the descending chain condition on right ideals, possesses a non-trivial commuting automorphism, then A is a field.

Fundamental to our discussion is the following:

LEMMA. *If T is a commuting automorphism of an associative ring A, then for every y and z of A we have*

$$(y - y^T) \cdot (zy - yz) = 0. \tag{1}$$

Proof. From $(x + y) \cdot (x + y)^T = (x + y)^T \cdot (x + y)$ we obtain

$$x \cdot y^T + y \cdot x^T = x^T \cdot y + y^T \cdot x$$

or in more convenient form

$$x \cdot y^T - y^T \cdot x = x^T \cdot y - y \cdot x^T. \tag{2}$$

From $(yx + y) \cdot (yx + y)^T = (yx + y)^T \cdot (yx + y)$ we obtain

$$y \cdot x \cdot y^T + y \cdot y^T \cdot x^T = y^T \cdot x^T \cdot y + y^T \cdot y \cdot x,$$

and since $y \cdot y^T = y^T \cdot y$ this can be written as

$$y \cdot (x \cdot y^T - y^T \cdot x) = y^T \cdot (x^T \cdot y - y \cdot x^T).$$

Using (2) this equation becomes

$$(y - y^T) \cdot (x^T \cdot y - y \cdot x^T) = 0.$$

Letting x^T be z and noting that z may be any element in A since T is an automorphism, we obtain (1) as desired.

The first result that comes from this Lemma, does not require any chain conditions.

[1]This paper was a small part of a Ph.D. thesis written under Professor A. A. Albert at the University of Chicago, 1950.

THEOREM 1. *If A is an associative division ring with a non-trivial commuting automorphism T, then A is a field.*

Proof. From (1) it is clear that every element y is either in the centre, or if it is not in the centre, then $y^T = y$. If A is not a field, take y not in the centre. Then $y^T = y$. Let c be any element in the centre. Then $y + c$ is not in the centre and therefore $(y + c)^T = y + c$. Thus $(y + c)^T = y^T + c^T = y + c^T = y + c$. Therefore $c^T = c$, $T = I$. This contradicts the non-triviality of T and therefore A is a field.

We shall now impose the condition that A has the descending chain condition on right ideals. If A is also a simple ring then it is a total matric ring with elements in a division ring.[2] We can now prove

THEOREM 2. *If A is a simple associative ring with descending chain condition on right ideals, and if A possesses a non-trivial commuting automorphism, then A is a field.*

Proof. If the matric ring has dimension 1, then the result follows immediately from Theorem 1. We therefore assume that the matric ring has dimension >1, and that it has a basis e_{ij} with the usual matric multiplication table. Every element in A is then a linear combination of the e_{ij} with coefficients in a division ring D.

Then

$$e_{ij}^T = \sum_{m,n} d_{mn} e_{mn},$$

with the d_{mn} in D. Using (1) we first let $y = e_{ij}$ and $z = e_{ki}$, with $k \neq j$. Then

$$\begin{aligned} 0 &= (e_{ij} - e_{ij}^T)\cdot(e_{ki}e_{ij} - e_{ij}e_{ki}) \\ &= - e_{ij}^T \cdot e_{kj} \\ &= - \sum_{m,n} d_{mn}e_{mn}\cdot e_{kj} \\ &= - \sum_{m} d_{mk}e_{mj}. \end{aligned}$$

Therefore $d_{mk} = 0$ for every $k \neq j$, and therefore

$$e_{ij}^T = \sum_m d_{mj}e_{mj}.$$

We consider now two cases, namely $j \neq i$ and $j = i$.

Case (a): $j \neq i$. Using (1) again, we let $y = e_{ij}$ and $z = e_{ji}$, and obtain

$$\begin{aligned} 0 &= (e_{ij} - e_{ij}^T)\cdot(e_{ji}e_{ij} - e_{ij}e_{ji}) \\ &= (e_{ij} - e_{ij}^T)\cdot(e_{jj} - e_{ii}) \\ &= (e_{ij} - \sum_m d_{mj}e_{mj})\cdot(e_{jj} - e_{ii}) \\ &= e_{ij} - \sum_m d_{mj}e_{mj}. \end{aligned}$$

[2]See for example *Rings with Minimum Condition* by Artin, Nesbitt, and Thrall.

Therefore

$$e_{ij} = \sum_m d_{mj} e_{mj} = e^T_{ij}.$$

Case (b): $j = i$. Using (1) we set $y = e_{ii}$ and $z = e_{ik}$, with $k \neq i$, and obtain

$$\begin{aligned} 0 &= (e_{ii} - e^T_{ii})\cdot(e_{ik}e_{ii} - e_{ii}e_{ik}) \\ &= - e_{ik} + e^T_{ii}\cdot e_{ik} \\ &= - e_{ik} + \sum_m d_{mi} e_{mi} e_{ik}. \end{aligned}$$

Therefore

$$e_{ik} = \sum_m d_{mi} e_{mk},$$

and thus for $m \neq i, d_{mi} = 0$, whereas for $m = i, d_{ii} = 1$. Therefore

$$e^T_{ii} = \sum_m d_{mi} e_{mi} = e_{ii}.$$

We have therefore shown that

$$e^T_{ij} = e_{ij}$$

for every i and j.

Now let $d \neq 0$ be any element of D. Again using (1) we set $y = d(e_{ij} + e_{ji})$ and $z = e_{jj}$, with $i \neq j$, and obtain

$$\begin{aligned} 0 &= [d(e_{ij} + e_{ji}) - d^T(e_{ij} + e_{ji})]\cdot[de_{ji} - de_{ij}] \\ &= d(d - d^T)(e_{ij} + e_{ji})(e_{ji} - e_{ij}) \\ &= d(d-d^T)(e_{ii} - e_{jj}). \end{aligned}$$

Therefore $d^T = d$, and $T = I$. This contradicts the non-triviality of T and therefore A is a field.

Finally we consider the consequences of a semi-simple ring with descending chain condition on right ideals, possessing a commuting automorphism.

THEOREM 3. *If A is a semi-simple associative ring with the descending chain condition on right ideals, and if A possesses a commuting automorphism, then A is a direct sum*

$$A = F_1 + F_2 + \ldots + F_m + B_1 + B_2 + \ldots + B_q$$

where the F_i are fields, where the B_j are simple ideals which are not fields and the contraction T_j of T onto B_j is the identity automorphism.

Proof. Since A is semi-simple, it is a direct sum of simple ideals[2]

$$A = A_1 + \ldots + A_n.$$

Since T is an automorphism, A_i^T is a simple ideal of A and is therefore one of the A_i. That is, T merely permutes the A_i.

We shall consider two cases, namely $A_i^T = A_i$, and $A_i^T = A_j$, with $j \neq i$.

Case (a). If $A_i^T = A_i$, then the contraction T_i of T onto A_i is a commuting automorphism of A_i. Therefore, since A_i is simple, either $T_i = I$, or if T_i is non-trivial, by Theorem 2, A_i is a field.

Case (b). If $A_i^T = A_j$, with $j \neq i$, then consider A_j^T. If $A_j^T = A_j$, the unity element e_j of A_j maps into itself, $e_j^T = e_j$. However since $A_i^T = A_j$, $e_i^T = e_j$ and therefore $(e_i - e_j)^T = 0$. Since $e_i - e_j \neq 0$, this is impossible and therefore $A_j^T = A_k$, with $k \neq j$. Of course k may be equal to i. Let x_i and x_j be any elements of A_i and A_j respectively, and consider their sum $x_i + x_j$. Since $(x_i + x_j)\cdot(x_i + x_j)^T = (x_i + x_j)^T\cdot(x_i + x_j)$ and since $A_i\cdot A_j = 0 = A_j\cdot A_j^T$ we obtain

$$x_i\cdot x_j^T + x_j\cdot x_i^T = x_i^T\cdot x_j + x_j^T\cdot x_i.$$

The products $x_j\cdot x_i^T$ and $x_i^T\cdot x_j$ are in A_j while the other two products are not in A_j. Therefore

$$x_j\cdot x_i^T = x_i^T\cdot x_j.$$

Therefore x_i^T is in the centre of A_j and therefore x_i is in the centre of A_i. Therefore A_i is commutative, A_i is a field. Q.E.D.

Thus a non-trivial commuting automorphism on a semi-simple associative ring with the descending chain condition on right ideals, leaves fixed all elements that are not in the centre. To see that non-central elements may be moved by a non-trivial commuting automorphism on a ring with a non-zero radical, we consider the following example.[3]

Let N be the ring generated by x and y such that $xy \neq yx$ and such that $N^3 = 0$. Let A be the ring obtained from N by adjoining a unity element to N. Then N is the radical of A, A is not semi-simple and not commutative. Consider the mapping T which sends a into $a^T = (1 - x + x^2)\cdot a\cdot(1 + x)$. Since $(1 + x)\cdot(1 - x + x^2) = 1$, T is an inner automorphism. Since $a = c_1 + c_2\cdot x + c_3\cdot y + c_4\cdot xy + c_5\cdot yx + c_6\cdot x^2 + c_7\cdot y^2$, c_i integers, a simple computation reveals that

$$a^T = a + c_3\cdot(yx - xy).$$

Thus it is clear that $a\cdot a^T = a^2 + c_1c_3(yx - xy) = a^T\cdot a$, and therefore T is a non-trivial commuting automorphism which moves y, an element not in the centre.

[3]This example was suggested to the author by Professor Jennings.

TRANSACTIONS OF THE ROYAL SOCIETY OF CANADA
VOLUME XLIX : SERIES III : JUNE, 1955
SECTION THREE

On a Theorem of Frobenius-König and J. von Neumann's Game of Hide and Seek[1]

LLOYD DULMAGE and ISRAEL HALPERIN, F.R.S.C.

1. Introduction. J. von Neumann **(9)** has introduced a game of hide and seek in connection with the optimal assignment problem. In this game, player I (the hider) selects one of the places in an $n \times n$ matrix of positive real α_{ij} $(i, j = 1, 2, \ldots, n)$ and player II (the seeker) selects a row or column of the matrix. If the seeker "finds" the hider in the row or column selected, the seeker receives α_{ij}. Otherwise he receives zero. In his paper, von Neumann has found optimal strategies for the hider.

The purpose of this note is to present a brief elementary proof of a theorem of G. Frobenius **(3)** and D. König **(7)** and to show its connection with this game of hide and seek. The results of section 2 and section 3 are not new. However it is hoped that our proof of the Frobenius-König theorem in section 2, and the notion of contemptuous strategies used in section 3 in finding optimal strategies for the hider, may be useful in discussing higher dimensional analogues of the game. In section 4, we use the Frobenius-König theorem to discuss optimal strategies for the seeker.

2. A theorem of Frobenius-König. A set of n places in an $n \times n$ matrix such that there is exactly one in each row and column, we call a *permutation set* of places. There are $n!$ such distinct sets. The Frobenius-König theorem asserts that a set S of places in an $n \times n$ matrix contains at least one place in every permutation set if and only if S contains an $s \times t$ submatrix of places, with $s + t = n + 1$. If the set S contains such a submatrix, it is immediate that it contains a place in every permutation set. For the necessity, we proceed by induction on n. The theorem is true for $n = 1$. If S contains all n^2 places the theorem is true. If not, consider the minor of a place w not in S. Every permutation set containing w contains an element of S, and, since w is not in S, every permutation set in the $(n-1) \times (n-1)$ minor contains a place in S. By the induction, the minor contains a $p \times q$ submatrix of places of S, with $p + q = (n-1) + 1 = n$. Consider the two complementary submatrices of dimensions $p \times p$ and $q \times q$. If both of these submatrices contain a permutation set of places not in S, then combining these we have such a permutation set for the whole matrix. Hence at least one of these submatrices, say the $p \times p$ submatrix, must contain a $u \times v$ submatrix of places in S with $u + v = p + 1$. Combining the $p \times q$ and $u \times v$ submatrices, we have a $(q + u) \times v$ submatrix of elements of S, and $q + u + v = q + p + 1 = n + 1$, as required.

[1]The research resulting in this note was carried out at the Summer Research Institute of the Canadian Mathematical Congress.

The Frobenius-König theorem is equivalent to the finite marriage problem as solved by P. Hall **(4)**, by H. Weyl **(12)**, by C. J. Everett and G. Whaples **(2)** and by P. R. Halmos and H. E. Vaughan **(5)**. The marriage problem formulation grew out of H. Weyl's consideration of Maak's Lemma **(8)**.

As a first corollary of this theorem we have the following result proved by König **(7)**. An $n \times n$ matrix of non-negative real numbers with row and column sums equal to $k > 0$ contains a permutation set of non-zero elements. For let S be the set of places in which the matrix elements are zero. If this set contains an $s \times t$ block of zeros with $s + t = n + 1$, the sum of the elements in the rows and columns intersecting in this block is $sk + tk = (n + 1)k$. This sum is less than or equal to the sum of all the elements, which is nk. Accordingly, by the Frobenius-König theorem there exists a permutation set of non-zero elements. If the elements of the matrix are restricted to non-negative integral values, we have, in this corollary, a proof of the existence of a complete system of representatives for any two divisions of mn things into m classes with n things in each class. This result has been proved by D. König **(7)**, B. L. Van der Waerden **(11)**, and E. Sperner **(10)**.

As a second corollary, we have the following result proved by A. J. Hoffman and H. W. Wielandt **(6)**, G. Birkhoff **(1)**, and J. von Neumann **(9)**. An $n \times n$ matrix which has all its elements zero except for a permutation set of ones is called a *permutation matrix*. An $n \times n$ matrix of non-negative real numbers in which all the row and column sums are 1 is called *doubly stochastic*. Our second corollary asserts that every doubly stochastic matrix is an average of some $n^2 - n + 1$ or less of the $n!$ permutation matrices. It follows that the set of doubly stochastic matrices, considered as a set of points in a space of n^2 dimensions, constitute the convex hull of the permutation matrices. To prove the second corollary, note that the first corollary provides us, in any doubly stochastic matrix, with a permutation set of positive elements. If $\lambda_1 > 0$ is the least of these elements and if A_1 is the permutation matrix with 1's in this permutation set, then $A - \lambda_1 A_1$ is an $n \times n$ matrix of non-negative elements which has all its row and column sums equal to $1 - \lambda_1$ and has at least one zero element. Applying the first corollary, in the same manner, to $A - \lambda_1 A_1$ we get a matrix $A - \lambda_1 A_1 - \lambda_2 A_2$ of non-negative elements with all its row and column sums equal to $1 - \lambda_1 - \lambda_2$ and at least two zero elements. After $k \leqslant n^2 - n$ such steps we arrive at a matrix of non-negative elements with exactly n elements >0 and with its row and column sums all equal to $1 - \lambda_1 - \ldots - \lambda_k$. These n elements must constitute a permutation set of equal elements. Hence we have $\lambda_i > 0$ and permutation matrices A_i such that

$$A = \sum_{i=1}^{k+1} \lambda_i A_i, \sum_{i=1}^{k+1} \lambda_i = 1,$$

and $k + 1 \leqslant n^2 - n + 1$.

3. Optimal strategies for the hider. In a zero-sum two person game, we shall call a player's strategy *contemptuous* if, for it, the payoff is independent of the opponent's strategy. In von Neumann's hide and seek game, every optimal strategy for the hider is contemptuous. To establish this, consider any strategy

$$(x_{ij}),\ x_{ij} \geqslant 0,\ \sum_{i,j} x_{ij} = 1, \qquad (i, j = 1, 2, \ldots, n),$$

and the matrix $A = (x_{ij}\,\alpha_{ij})$. The row and column sums of A are the payoffs for the hider's strategy (x_{ij}) corresponding to the seeker's pure strategies. Let K be the maximum of these. If the strategy (x_{ij}) is not contemptuous, at least one row or column sum is $<K$. Since the sum of the row sums equals the sum of the column sums, at least one row and at least one column has sum $<K$. Let the common element be $x_{pq}\,\alpha_{pq}$. If we let

$$x'_{ij} = \begin{cases} \dfrac{x_{ij} + \epsilon}{1 + \epsilon} & \text{for } i = p, j = q, \\[2ex] \dfrac{x_{ij}}{1 + \epsilon} & \text{otherwise,} \end{cases}$$

we have a strategy

$$x'_{ij} \geqslant 0,\ \sum_{i,j} x'_{ij} = 1,$$

and the row and column sums of $A' = (x_{ij}'\,\alpha_{ij})$ are less than the corresponding sums of A, and hence less than K, except possibly for the qth row and pth column. If we choose

$$\epsilon = \frac{K - \max\left[\sum_i x_{iq},\ \sum_j x_{pj}\right]}{2\,\alpha_{pq}},$$

these too are $<K$. Accordingly, the maximum of the row and column sums of A' is $<K$ and (x_{ij}) is not an optimal strategy.

Let i^p denote the mate of i in any permutation P of $1, 2, \ldots, n$. If δ_{ij} is the Kroneker delta, the permutation matrix corresponding to the permutation P has elements $a_{ij} = \delta_{i^p, j}$. For any contemptuous strategy (x_{ij}), denote the equal row and column sums of $A = (x_{ij}\,\alpha_{ij})$ by T. Optimal strategies are those contemptuous strategies for which T is least. From the second corollary of section 2, we see that, for any contemptuous strategy (x_{ij}), there exist T_p such that the matrix $(x_{ij}\,\alpha_{ij}) = \sum_p T_p \delta_{(i^p, j)}$, $T_p \geqslant 0$, $\sum_p T_p = T$, there being a term (possibly zero) in these summations for each of the $n!$ permutations. Since $\sum_{ij} x_{ij} = 1$, we have

$$\sum_p T_p\left(\sum_i \frac{1}{\alpha_{i,\,i^p}}\right) = 1.$$

The optimal strategies for the hider are those which minimize $T = \sum_p T_p$ subject to the constraints

$$T_p \geqslant 0 \text{ and } \sum_p T_p\left(\sum_i \frac{1}{\alpha_{i,\,i^p}}\right) = 1.$$

Let

$$M = \max_{P} \sum_{i} \frac{1}{\alpha_{i,\,i^P}}.$$

We minimize T by taking $T_P = 0$ unless

$$\sum_{i} \frac{1}{\alpha_{i,\,i^P}} = M.$$

The minimum T, which is the value of the game, is $1/M$. If there is a unique permutation P' for which

$$M = \sum_{i} \frac{1}{\alpha_{i,\,i^{P'}}},$$

then there is a unique optimal strategy

$$x_{ij} = \begin{cases} \dfrac{1}{M\alpha_{ij}} & \text{for } j = i^{P'}. \\ 0, & \text{otherwise.} \end{cases}$$

If there is more than one permutation P such that

$$M = \sum_{i} \frac{1}{\alpha_{i,\,i^P}},$$

then each such permutation yields a basic optimal strategy of this form and any other optimal strategy is an average of these.

4. Optimal strategies for the seeker. The results in this section could be found using section 3 and the fundamental theorem[2] on zero-sum two-person games. Since it is our purpose to show the direct connection between the Frobenius-König theorem and the hide and seek game, we proceed independently.

Let (r_i, c_j) $i, j = 1, 2, \ldots, n$, $r_i \geqslant 0$, $c_j \geqslant 0$, $\sum_i r_i + \sum_j c_j = 1$, be row and column probabilities which together constitute a strategy for the seeker. Denote $(r_i + c_j)\,\alpha_{ij}$ by K_{ij}. If the strategy is optimal, we have $K_{ij} \geqslant$ the value v of the game for all i, j and $K_{ij} = v$ for some i, j. Hence $(r_i + c_j) \geqslant v/\alpha_{ij}$ for all i, j. Since $\sum_i (r_i + c_{i^P}) = 1$ for all permutations P, we have

$$v \leqslant \frac{1}{\displaystyle\sum_{i} \frac{1}{\alpha_{i,\,i^P}}}$$

for all P. Considering any strategy in which $r_i, c_j > 0$ for all i, j we have $K_{ij} > 0$ for all i, j so that $v > 0$. Accordingly $0 < v \leqslant 1/M$ where

$$M = \max_{P} \sum_{i} \frac{1}{\alpha_{i,\,i^P}}.$$

[2]The fundamental theorem of J. von Neumann asserts that the value of the game to player I is equal to the value of the game to player II.

If $r_i = c_j = 0$ for some i, j, the strategy is not optimal, for the corresponding $K_{ij} = 0 < v$. Hence, in any optimal strategy, either $r_i > 0$ for all i or $c_j > 0$ for all j. If $r_i > 0$ for all i and if ϵ is less than the least of r_i, and if $r_i' = r_i - \epsilon$, $c_j' = c_j + \epsilon$, we have $r_i' > 0$, $c_j' > 0$, $\sum_i r_i' + \sum_j c_j' = 1$, and $K_{ij}' = K_{ij}$ for all i, j. Two such strategies as (r_i, c_j) and (r_i', c_j') which have the property that the expectations corresponding to each of the opponents' strategies are equal, we call *opponent-equivalent*. For every optimal strategy for the seeker, there corresponds a family of opponent-equivalent strategies. There are two extreme strategies of the family; in one of these some r_i is zero and, in the other, some c_j is zero. Every strategy of the family is an average of these two extremes.

Consider any optimal strategy for the seeker and let (r_i, c_j) denote an opponent-equivalent member of the corresponding family other than one of the two extremes, so that $r_i > 0$, $c_j > 0$ for all i, j. Let the set S of places in the matrix (K_{ij}) consist of all elements of (K_{ij}) which are $>v$. Suppose there exists an $s \times t$ submatrix of places of S with $s + t = n + 1$. For definiteness, let this submatrix consist of K_{ij} for $i = 1, 2, \ldots, s$ and $j = 1, 2, \ldots, t$. Consider

$$r'_i = \begin{cases} r_i - \epsilon \text{ for } i = 1, 2, \ldots, s, \\ r_i + \epsilon + \dfrac{2\epsilon}{n-1} \text{ for } i = s + 1, \ldots, n, \end{cases}$$

$$c'_j = \begin{cases} c_j - \epsilon \text{ for } j = 1, 2, \ldots t, \\ c_j + \epsilon + \dfrac{2\epsilon}{n-1} \text{ for } j = t + 1, \ldots, n, \end{cases}$$

in which $\epsilon > 0$ is small enough that $K'_{ij} = (r_i' + c_j')\,\alpha_{ij} > v$ for $i = 1, 2, \ldots, s$ and $j = 1, 2, \ldots, t$ and that $r_i', c_j' > 0$. This is ensured by choosing ϵ less than the least of $(K_{ij} - v)/2\alpha_{ij}$, r_i and c_j for $i = 1, \ldots, s$ and $j = 1, \ldots, t$. We have $\sum r_i' + \sum c_j' = 1$. For $i > s$ or $j > t$ we have $K'_{ij} > K_{ij}$. Hence

$$\min_{i,j} (K'_{ij}) > v,$$

so that the strategy (r_i, c_j) is not optimal. Accordingly, by the Frobenius-König theorem, the matrix (K_{ij}) contains at least one permutation set of places in which $K_{ij} = v$. There exists at least one permutation P' for which $K_{i,iP'} = v$, that is $r_i + c_{iP'} = v/\alpha_{i,iP'}$. Hence

$$v = \frac{1}{\sum_i \alpha_{i,iP'}} = \frac{1}{M}.$$

The maximum

$$\sum_i \frac{1}{\alpha_{i,i^P}}$$

is attained when $P = P'$.

By interchanging rows and columns in the game matrix (α_{ij}) and in the matrix (K_{ij}), we can arrange that the permutation P' is the identity. Hence for any optimal strategy (r_i, c_j) we have

$$r_i + c_i = \frac{v}{\alpha_{ii}} \text{ for } i = 1, 2, \ldots, n,$$

$$r_i + c_j \geqslant \frac{v}{\alpha_{ij}} \text{ for } i \neq j.$$

Hence for the c_i of an optimal strategy, we have the two sets of inequalities

$$\text{(1)} \qquad 0 \leqslant c_i \leqslant \frac{v}{\alpha_{ii}}, \text{ for } i = 1, 2, \ldots n,$$

and

$$\text{(2)} \qquad \frac{v}{\alpha_{ii}} - \frac{v}{\alpha_{ij}} \geqslant c_i - c_j \geqslant \frac{v}{\alpha_{ji}} - \frac{v}{\alpha_{jj}} \text{ for } i \neq j$$

Let $(c_1, c_2, \ldots, c_n)$ denote a point in a space of n-dimensions. Each of the $\binom{n}{2}$ inequalities (2) defines a region bounded by two (possibly coincident) parallel hyperplanes in which this point $(c_1, c_2, \ldots, c_n)$ must lie. The intersection of these $\binom{n}{2}$ regions is convex. Moreover, if $(c_1, c_2, \ldots, c_n)$ is in S then $(c_1 + k, c_2 + k, \ldots, c_n + k)$ is in S for arbitrary k. In other words, the inequalities (2) define a convex set of lines all of which are parallel to the vector

$$e = \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix}.$$

Any line m of this set contains a unique point $P(c_1', c_2', \ldots, c_n')$ in which at least one of the c_i' is zero and the others are positive. Using (2), we have $0 \leqslant c_i' < v/\alpha_{ii}$ for $i = 1, 2, \ldots, n$. If we denote the least of $v/\alpha_{ii} - c_i'$ by $k > 0$, then the point $Q(c_1' + k, c_2' + k, \ldots, c_n' + k)$ is on m and the points of the segments PQ have co-ordinates which satisfy the inequalities (1). The co-ordinates c_j of any point m of the segment PQ, together with $r_i = v/\alpha_{ii} - c_i$, yield an optimal strategy (r_i, c_j). The other points of the segment PQ, in the same way, yield strategies of the opponent-equivalent family of which (r_i, c_j) is a member. The points P and Q yield the extreme strategies of this family. Thus every line m of the intersection S provides us with a family of optimal opponent-equivalent strategies for the seeker.

References

1. G. Birkhoff, *Tres observaciones sobre el algebra lineal*, Rev. Universidad Nacional de Tucuman, Series A, *5* (1946), 147–150.
2. C. J. Everett and G. Whaples, *Representations of sequences of sets*, Amer. J. Math., *71* (1949), 287–293.
3. G. Frobenius, *Ueber Matrizen mit nicht negativen Elementen*, S.B. Berliner Akad., *23* (1912), 456–477.
4. P. Hall, *On representation of subsets*, J. London Math. Soc., *10* (1935), 26–30.
5. P. R. Halmos and H. E. Vaughan, *The marriage problem*, Amer. J. Math., *72* (1950), 214–215.
6. A. J. Hoffman and H. W. Wielandt, *The variation of the spectrum of a normal matrix*, Duke Math. J., *20* (1953), 37–39.
7. D. König, *Ueber Graphen und ihre Anwendung*, Math. Ann., *77* (1916), 453–465.
8. W. Maak, *Eine neue Definition der fastperiodischen Funktionen*, Abh. Math. Sem. Hamburgischen Univ., *11* (1936), 240–244.
9. J. von Neumann, *A certain zero-sum two-person game equivalent to the optimal assignment problem*, Contributions to the Theory of Games, vol. II, Ann. Math. Studies, no. 28, pp. 5–12.
10. E. Sperner, *Note zu der Arbeit von Herrn B.L. van der Waerden: Ein Satz über Klasseneinteilungen von endlichen Mengen*, Abh. Math. Sem. Hamburgischen Univ., *5* (1926), 232.
11. B. L. van der Waerden, *Ein Satz über Klasseneinteilungen von endlichen Mengen*, Abh. Math. Sem. Hamburgischen Univ., *5* (1926), 185–188.
12. H. Weyl, *Almost periodic invariant vector sets in a metric vector space*, Amer. J. Math., *71* (1949), 178–205.

TRANSACTIONS OF THE ROYAL SOCIETY OF CANADA
VOLUME XLIX : SERIES III : JUNE, 1955
SECTION THREE

Radical Rings with Nilpotent Associated Groups[1]

S. A. JENNINGS

Presented by R. D. JAMES, F.R.S.C.

1. Introduction. An associative ring is said to be a *radical ring* (in sense of Jacobson **(2)**) if for every element x in the ring there exists in the ring a "quasi-inverse" x' such that

$$x + x' + xx' = 0. \tag{1.0.1}$$

Every nil-ring (and *a fortiori* every nilpotent ring) is a radical ring, but there are radical rings which are not nil.

Let R be a radical ring. Then the elements of R form a group $\mathfrak{G}$ under the operation $*$ of quasi-multiplication, defined by

$$x*y = x + y + xy \tag{1.0.2}$$

for all $x, y \in R$, the element 0 of R being the unit element of $\mathfrak{G}$ and the element x' being the inverse, relative to the operation $*$, of x. Indeed, the fact that the elements form a group under $*$ is characteristic of radical rings. We say that $\mathfrak{G}$ is the *group associated with* R, and when necessary write $\mathfrak{G} = \mathfrak{G}\ (R)$ to stress this association. We consider also the Lie ring Λ associated with R, where Λ consists of the elements of R under addition and commutation, where as usual for $x, y \in R$ the Lie commutator $x \circ y$ is defined by setting $x \circ y = xy - yx$.

In this note we investigate some of the relationships between the structures of R, $\mathfrak{G}$ and Λ, and in particular we prove that if $\mathfrak{G}$ is a nilpotent group, then Λ is a nilpotent Lie ring, and conversely.

2. Commutators in G and Λ. A radical ring cannot have a 1-element, since the existence of $(-1)'$ would imply $-1 = 0$ by (1.0.1). In what follows we find it convenient to adjoin in a formal manner to the radical ring R an element 1 having the property $x1 = 1x = x$ for all $x \in R$. We will then in considering $\mathfrak{G}$ replace the $*$-multiplication (1.0.2) by formal multiplication of elements of the form $X = 1 + x$, $Y = 1 + y$, viz.,

$$XY = (1 + x)(1 + y) = 1 + x + y + xy$$

so that $XY = 1 + x*y$, and, if we define X^{-1} as $1 + x'$, we have $XX^{-1} = 1 + x + x' + xx' = 1 + 0 = 1$. No inconsistencies are introduced, there-

[1]This paper was written while the author was a Fellow at the Summer Research Institute of the Canadian Mathematical Congress, 1955. The author takes this opportunity of expressing his gratitude to the Canadian Mathematical Congress and the National Research Council.

fore, if we consider the group $\mathfrak{G}$ associated with R to be the multiplicative group of elements of the form $(1 + x)$. We note that expressions of the form $X + y$, Xy, yX, $X - Y$, etc. are defined either in $\mathfrak{G}$ or in R, as the case may be, but that, for example, $X + Y$ is meaningless for us. We note, too, that if $X = 1 + x$ and $x, y, z \in R$, then

$$y + yx = yX = z$$

implies

$$y = zX^{-1} = z + zx'$$

and in particular

$$Xy = 0 \text{ implies } y = 0 \tag{2.0.1}$$
$$yX = 0 \text{ implies } y = 0.$$

We use lower case letters for elements of R, and as far as is feasible indicate elements of $\mathfrak{G}$ by the corresponding capital, i.e., if $a \in R$ we write $A = 1 + a$, etc.

Consider the element $(X, Y) = X^{-1}Y^{-1}XY$ in $\mathfrak{G}$: if $X = 1 + x$, $Y = (1 + y)$, we have

$$(X, Y) = 1 + X^{-1}Y^{-1}(XY - YX)$$
$$= 1 + X^{-1}Y^{-1}(x \circ y) \tag{2.0.2}$$

and if $Z = 1 + z = (X, Y)$, then

$$(x \circ y) = YXz = z + yz + xz + yxz. \tag{2.0.3}$$

Formulae (2.0.2) and (2.0.3) reveal the intimate connection between commutation in $\mathfrak{G}$ and in Λ. In what follows we denote the ith subgroup of the lower central series of $\mathfrak{G}$ by $\mathfrak{G}_i$, where $\mathfrak{G}_i$ is defined inductively by

$$\mathfrak{G}_1 = \mathfrak{G}, \quad \mathfrak{G}_i = (\mathfrak{G}_{i-1}, \mathfrak{G})$$

as usual, and we denote the ρth term of the lower central series of Λ by Λ^ρ, and again Λ^ρ is defined inductively by

$$\Lambda = \Lambda^1, \quad \Lambda^\rho = \Lambda^{\rho-1} \circ \Lambda,$$

where if Γ, Δ are two ideals of Λ we mean by $\Gamma \circ \Delta$ the ideal spanned by all elements of the form $c \circ d$, $c \in \Gamma$, $d \in \Delta$. If for $\mathfrak{G}$ we have

$$\mathfrak{G} = \mathfrak{G}_1 \supset \mathfrak{G}_2 \supset \ldots \supset \mathfrak{G}_c \supset \mathfrak{G}_{c+1} = (1) \tag{2.0.4}$$

then $\mathfrak{G}$ is said to be nilpotent of class c, while if

$$\Lambda = \Lambda^1 \supset \Lambda^2 \supset \ldots \supset \Lambda^\gamma \supset \Lambda^{\gamma+1} = (0)$$

then Λ is nilpotent of class γ.

We have at once the following

LEMMA 2.1. *The centres of $\mathfrak{G}$, Λ and R consist of the same elements.*

Proof. If c is in the centre of R, i.e. if $cx = xc$ for all $x \in R$, then clearly $CX = XC$ where $C = 1 + c$, $X = 1 + x$, and $c \circ x = cx - xc = 0$, and conversely,

LEMMA 2.2. *If $\mathfrak{G}$ is nilpotent of class c, and $H_c = 1 + h_c \in \mathfrak{G}_c$, then h_c is in the centre of R. In particular, if $H_{c-1} = 1 + h_{c-1} \in \mathfrak{G}_{c-1}$ and $X = 1 + x$, then*

$$H_{c-1}^{-1}X^{-1}(h_{c-1} \circ x),$$

is in the centre of R for all elements $x \in R$.

Proof. Since

$$1 + H_{c-1}^{-1}X^{-1}(h_{c-1} \circ x) \in \mathfrak{G}_c \text{ if } 1 + h_{c-1} \in \mathfrak{G}_{c-1}$$

by (2.0.2), the lemma follows at once from the remark that if $\mathfrak{G}$ is nilpotent then $\mathfrak{G}_c$ is contained in the centre of $\mathfrak{G}$.

LEMMA 2.3. *For all $a, x, y, z \in R$ we have*[2]

$$[X(y \circ z)] \circ a = X(y \circ z \circ a) + (x \circ a)(y \circ z)$$

(where as usual $X(y \circ z)$ will denote $(y \circ z) + x(y \circ z)$). In particular, if a commutes with x

$$[X(y \circ z)] \circ a = X(y \circ z \circ a).$$

Proof. The result follows from the identity $(bc) \circ a = b(c \circ a) + (b \circ a)c$.

For later use we note also the following identities, which may be deduced from (1.0.1),

$$(2.3.1) \qquad \begin{aligned} (a \circ x)X^{-1} &= -X(a \circ x'), \\ X^{-1}(a \circ x) &= -(a \circ x')X. \end{aligned}$$

3. The principal theorem. We are now in a position to prove the following:

THEOREM 3.1. *If the group $\mathfrak{G}$ associated with a radical ring R is nilpotent, then the Lie ring Λ associated with R is also nilpotent.*

Proof. If $\mathfrak{G}$ is commutative, then so is R, and $\Lambda^2 = 0$, by (2.1), so that the theorem is trivially true when the class c of $\mathfrak{G}$ is 1. We suppose, therefore, that $\mathfrak{G}$ is nilpotent of class $c > 1$, and that the lower central series of $\mathfrak{G}$ is as in (2.0.4). Let $H_{c-1} = 1 + h_{c-1}$ be an element of $\mathfrak{G}_{c-1}$: then by (2.2),

$$H_{c-1}^{-1}X^{-1}(h_{c-1} \circ x)$$

is in the centre of R, so that, for all $y \in R$

$$(3.1.1) \qquad [H_{c-1}^{-1}X^{-1}(h_{c-1} \circ x)] \circ y = 0.$$

[2]Throughout this paper we write $((a \circ b) \circ c) = a \circ b \circ c$, $(((a \circ b) \circ c) \circ d) = a \circ b \circ c \circ d$, etc.

In particular, by (2.3) and the fact that $XH_{c-1} = 1 + x + h_{c-1} + xh_{c-1}$ commutes with $H_{c-1}{}^{-1} X^{-1}$,

$$[H_{c-1}^{-1}X^{-1}(h_{c-1} \circ x)] \circ (x + h_{c-1} + xh_{c-1})$$
$$= H_{c-1}^{-1}X^{-1}[h_{c-1} \circ x \circ (x + h_{c-1} + xh_{c-1})] = 0$$

so that, by (2.0.1)

(3.1.2) $$h_{c-1} \circ x \circ (x + h_{c-1} + xh_{c-1}) = 0$$

for all $x \in R$ and all h_{c-1} such that $1 + h_{c-1} \in \mathfrak{G}_{c-1}$. Set $x = a + h_{c-1}'$ in (3.1.2), where a is any element of R: then (3.1.2) becomes

(3.1.3) $$h_{c-1} \circ (a + h'_{c-1}) \circ (a + h'_{c-1} + h_{c-1} + ah_{c-1} + h'_{c-1}h_{c-1}) = 0$$

and since, for any element b in R, $bb' = b'b$, we have, from (3.1.3) and (1.0.1)

(3.1.4) $$h_{c-1} \circ a \circ (a + ah_{c-1}) = 0.$$

However, (3.1.2) with $x = a$ takes the form

(3.1.5) $$h_{c-1} \circ a \circ (a + h_{c-1} + ah_{c-1}),$$

so that, subtracting (3.1.4) from (3.1.5), we get

(3.1.6) $$h_{c-1} \circ a \circ h_{c-1} = 0.$$

Now we may write (3.1.4) in the form

$$(h_{c-1} \circ a \circ aH_{c-1}) = 0$$

and hence

$$(h_{c-1} \circ a \circ a)H_{c-1} + a(h_{c-1} \circ a \circ h_{c-1}) = 0$$

so that, by (3.1.6) and (2.0.1),

(3.1.7) $$h_{c-1} \circ a \circ a$$

for all $a \in R$, and $H_{c-1} = 1 + h_{c-1} \in \mathfrak{G}_{c-1}$.

From (3.1.7) with $a = y + z$ we get,

(3.1.8) $$(h_{c-1} \circ y \circ z) + (h_{c-1} \circ z \circ y) = 0$$

and in (3.1.8) using $y = uv, z = v$ we get

$$h_{c-1} \circ uv \circ v + (h_{c-1} \circ v) \circ (uv) = 0$$
$$= [(h_{c-1} \circ u)v + u(h_{c-1} \circ v)] \circ v + (h_{c-1} \circ v \circ u)v + u(h_{c-1} \circ v \circ v)$$
$$= (h_{c-1} \circ u \circ v)v + (h_{c-1} \circ v \circ u) + (u \circ v)(h_{c-1} \circ v)$$

so that from (3.1.8.) we have, using (3.1.7),

(3.1.9) $$(u \circ v)(h_{c-1} \circ v) = 0.$$

Similarly with $y = vu$ and $z = v$ in (3.1.8) we may obtain

(3.1.10) $$(h_{c-1} \circ v)(u \circ v) = 0$$

for all $u, v \in R$, and $H_{c-1} = 1 + h_{c-1} \in \mathfrak{G}_{c-1}$.

From (3.1.8) again with $y = h_{c-1}u$ and $z = v$ we have

$$\begin{aligned} h_{c-1} \circ (h_{c-1}u) \circ v + h_{c-1} \circ v \circ (h_{c-1}u) &= 0 \\ = [h_{c-1}(h_{c-1} \circ u)] \circ v &+ (h_{c-1} \circ v \circ h_{c-1})u + h_{c-1}(h_{c-1} \circ v \circ u) \\ = (h_{c-1} \circ v)(h_{c-1} \circ u) &+ h_{c-1}(h_{c-1} \circ u \circ v) + h_{c-1}(h_{c-1} \circ v \circ u) \end{aligned}$$

so that, using (3.1.8) again,

(3.1.11) $$(h_{c-1} \circ v)(h_{c-1} \circ u) = 0.$$

Returning to (3.1.1), we may re-write it, using (2.3), in the form

(3.1.12) $$0 = H_{c-1}^{-1}X^{-1}(h_{c-1} \circ x \circ y) + H_{c-1}^{-1}(x' \circ y)(h_{c-1} \circ x) + (h'_{c-1} \circ y)X^{-1}(h_{c-1} \circ x).$$

Now consider, bearing in mind that $(x \circ y) = -(y \circ x)$, the second term on the right in (3.1.12). We have

$$\begin{aligned} H_{c-1}^{-1}(x' \circ y)(h_{c-1} \circ x)X &= H_{c-1}^{-1}(x' \circ y)X(h_{c-1} \circ x) && \text{by (3.1.7)} \\ &= H_{c-1}^{-1}X^{-1}(x \circ y)(h_{c-1} \circ x) && \text{by (2.3.1)} \\ &= 0 && \text{by (3.1.9)} \end{aligned}$$

so that $H_{c-1}^{-1}(x' \circ y)(h_{c-1} \circ x) = 0$ by (2.0.1). Again, consider the third term on the right of (3.1.12). We have

$$(h_{c-1} \circ x)XH_{c-1} = XH_{c-1}(h_{c-1} \circ x) \quad \text{by (3.1.6) and (3.1.7)}$$

so that,

$$\begin{aligned} (h'_{c-1} \circ y)X^{-1}(h_{c-1} \circ x)XH_{c-1} &= (h'_{c-1} \circ y)H_{c-1}(h_{c-1} \circ x) \\ &= -H_{c-1}^{-1}(h_{c-1} \circ y)(h_{c-1} \circ x) && \text{by (2.3.1)} \\ &= 0 && \text{by (3.1.11)} \end{aligned}$$

so that $(h_{c-1}' \circ y)X^{-1}(h_{c-1} \circ x) = 0$ and in (3.1.12) we have

$$H_{c-1}X^{-1}(h_{c-1} \circ x \circ y) = 0$$

and finally, by (2.3.1),

(3.1.13) $$(h_{c-1} \circ x \circ y) = 0.$$

Because of (3.1.13) and the Jacobi identity we have

$$h_{c-1} \circ (u \circ v) = 0, \qquad u, v \in R,$$

so that, since $h_{c-1} \circ [x(u \circ v)] \circ w = 0$ by (3.1.13), we obtain

(3.1.14) $$(h_{c-1} \circ x)(u \circ v \circ w) = 0.$$

Having established (3.1.13) and (3.1.14), Theorem (3.1) may be proved as follows: Consider the ideal S of R generated by the elements $(h_{c-1} \circ x)$,

where $1 + h_{c-1} \in \mathfrak{G}_{c-1}$. Every element of $\mathfrak{G}_c$ may be written as a product of elements of the form $1 + h_c$, where

$$h_c = X^{-1}H_{c-1}^{-1}(h_{c-1} \circ x)$$

so that S contains all elements z such that $1 + z \in \mathfrak{G}_c$. Every element s of S is expressible as sums of elements of the form $h_{c-1} \circ x$ and of the form $u(h_{c-1} \circ x)$ where $x, u \in R$, because of (3.1.13). Consider $\bar{R} = R/S$: $\bar{R}$ is a radical ring, and $\mathfrak{G}(\bar{R})$ is homomorphic to $\mathfrak{G}(R)$ under the mapping $R \to \bar{R}$:

$$x \to \bar{x} = x + S,$$

but since $z \in S$ when $1 + z \in \mathfrak{G}_c$, $\mathfrak{G}(\bar{R})$ is nilpotent of class less than c. Assume that (3.1) is established for rings whose groups are of class less than c, so that in particular $\Lambda(\bar{R})$ is a nilpotent Lie ring of class say $\bar{\gamma}$. Then since $\Lambda(R) \to \Lambda(\bar{R})$ under the mapping $R \to \bar{R}$ as above, any commutator of the form

$$c_1 \circ c_2 \circ \ldots \circ c_{\bar{\gamma}} \circ c_{\bar{\gamma}+1}$$

belongs to S. However, if $s \in S$ then

$$s \circ u \circ w = 0,$$

since any element of S is a sum of elements of the forms $h_{c-1} \circ x$ or $u(h_{c-1} \circ x)$, and by (3.1.13) $h_{c-1} \circ x \circ v = 0$, and $(u \circ v \circ w)(h_{c-1} \circ x) = 0$. In $\Lambda(R)$, therefore,

$$c_1 \circ c_2 \circ \ldots \circ c_{\bar{\gamma}+3} = 0$$

and $\Lambda(R)$ is of class $\gamma = \bar{\gamma} + 2$ at most.

We prove also the converse of (3.1), viz.

THEOREM 3.2. *If R is a radical ring with $\Lambda(R)$ a nilpotent Lie ring then $\mathfrak{G}(R)$ is a nilpotent group.*

Proof. Let Λ be of class $\gamma > 1$, since the result is trivial if $\gamma = 1$, and let $c_{\gamma-1} \in \Lambda^{\gamma-1}$. Then $c_{\gamma-1} \circ x \in \Lambda^{\gamma}$, and $c_{\gamma-1} \circ x \circ y = 0$ for all $x, y \in R$. Then by the Jacobi identity $c_{\gamma-1} \circ (x \circ y) = 0$ so that

$$\begin{aligned} c_{\gamma-1} \circ [x(u \circ v)] \circ y &= 0 \\ &= (c_{\gamma-1} \circ x \circ y)(u \circ v) + (c_{\gamma-1} \circ x)(u \circ v \circ y) \end{aligned}$$

and we have

(3.2.1) $$(c_{\gamma-1} \circ x)(u \circ v \circ y) = 0.$$

Similarly, since $c_{\gamma-1} \circ xy \circ x = 0$, we see that

(3.2.2) $$(c_{\gamma-1} \circ x)(y \circ x) = 0$$

Let T be the ideal of R generated by all elements of the form $(c_{\gamma-1} \circ x)$: every element in T may be written as a sum of elements of one or other of the forms $c_{\gamma-1} \circ x$, $u(c_{\gamma-1} \circ x)$ with $u \in R$. In particular the elements of Λ^{γ} belong to R, so that $\Lambda(R/T)$ is of class $<\gamma$. By an induction assumption we may assume $\mathfrak{G}(R/T)$ is nilpotent of class $\bar{c}$, say, and every element

$h_{\bar{c}+1}$, where $1 + h_{\bar{c}+1} \in \mathfrak{G}_{\bar{c}+1}$, belongs to T. We will verify that $\mathfrak{G}_{\bar{c}+3} = 0$ by showing that, if $g \in T$, then $(G, U, V,) = 1$, when $G = 1 + g$, $U = 1 + u$, $V = 1 + v$ for all $u, v \in R$. Indeed, if $(G, U) = 1 + c$ we have seen in (2.0.2) that

$$c = G^{-1}U^{-1}(g \circ u)$$

and so it will be sufficient to prove that $c \circ v = 0$ for all $v \in R$, and where g is a sum of elements of the form $x(c_{\gamma-1} \circ y)$, since we know $c_{\gamma-1} \circ y \circ x = 0$ already.

Now $c \circ v = G^{-1}U^{-1}(g \circ u \circ v) + G^{-1}(u' \circ v)(g \circ u) + (g' \circ v)U^{-1}(g \circ v)$ However,

$$[x(c_{\gamma-1} \circ y)] \circ u = (x \circ u)(c_{\gamma-1} \circ y)$$

and

$$\begin{aligned}[x(c_{\gamma-1} \circ y)] \circ u \circ v &= (x \circ u \circ v)(c_{\gamma-1} \circ y) \\ &= 0 \qquad \text{by (3.2.1)}\end{aligned}$$

so that, since g is a sum of elements of the form $x(c_{\gamma-1} \circ y)$

(3.2.3) $$g \circ u \circ v = 0.$$

Similarly

$$\begin{aligned}&(u' \circ v)(x \circ u)(c_{\gamma-1} \circ y) \\ &= [(u'x) \circ u \circ v - u'(x \circ v \circ u)](c_{\gamma-1} \circ y) \\ &= 0 \qquad \text{by (3.2.1)}\end{aligned}$$

so that $(u' \circ v)[x(c_{\gamma-1} \circ y) \circ v] = 0$, and hence $G^{-1}(u' \circ v)(g \circ u) = 0$. Finally

$$\begin{aligned}&(g' \circ v)U^{-1}(g \circ u) \\ &= (g' \circ v)(g \circ u)U^{-1} \qquad \text{by (3.2.3)} \\ &= [g'(u \circ v \circ g) - (g'u \circ g \circ v)]U^{-1} \\ &= 0, \qquad \text{by (3.2.3)}\end{aligned}$$

and we see that $c \circ v = 0$ for all $v \in R$.

The following now follows at once from the result in **(3)**:

COROLLARY 3.3. *If R is a radical ring whose associated group is nilpotent, then R contains a nil-ideal N such that R/N is commutative.*

4. Nilpotent algebras of characteristic zero. From the foregoing proofs of (3.1) and (3.2) it will be noted that if $c = 2$, we have $\gamma = 2$, since then h_{c-1} is any element of R, and equally, if $\gamma = 2$, then $c_{\gamma-1}$ is any element of R, $T = R$, and $\mathfrak{G}$ is of class 2. It seems reasonable to conjecture that c is always equal to γ, but we have not been able to prove this for radical rings in general. However, if R is a nil algebra over a field of characteristic zero, then we prove that $c = \gamma$, provided either c or γ (and hence both) are finite.

Let A be a nil-algebra over a field Φ of characteristic 0, that is, an algebra such that for all $x \in A$ there exists an integer ν, perhaps depending on x,

such that $x^\nu = 0$. It is well known that every such A is a radical ring, with

$$x' = -x + x^2 - x^3 + \ldots \pm x^{\nu-1}.$$

Since Φ is of characteristic 0, we may express every element $X = 1 + x$ in the form

$$X = (1 + x) = \exp(\xi)$$

where

$$\xi = \log(1 + x) = x - \frac{x^2}{2} + \frac{x^3}{3} - \ldots \frac{\pm x^{\nu-1}}{(\nu - 1)}$$

and where, since $\xi^\nu = 0$,

$$\exp \xi = 1 + \xi + \frac{\xi^2}{2!} + \ldots + \frac{\xi^{\nu-1}}{(\nu - 1)!}$$

Assume that Λ is of class γ. Then, for $X = 1 + x$, $Y = 1 + y$, by the Hausdorff-Campbell formula (cf. **1**, formula 26), we have

$$X^{-1}Y^{-1}XY = \exp(\xi \circ \eta + \sum c_i)$$

where $\eta = \log(1 + y)$, and $\sum c_i$ is a sum of commutators c_i in ξ and η of weight greater than 2, the sum being finite since Λ is of finite class. More generally,

$$(X_1, X_2, \ldots, X_n) = \exp(\xi_1 \circ \xi_2 \circ \ldots \circ \xi_n + \sum d_j)$$

where $\sum d_j$ is a finite sum of commutators d_j in $\xi_1 = \log X_1, \ldots, \xi_n = \log X_n$ of weight greater than n. In particular

$$(X_1, X_2, \ldots, X_\gamma) = \exp(\xi_1 \circ \xi_2 \circ \ldots \circ \xi_\gamma)$$

where all commutators of weight greater than γ are 0, since $\Lambda^{\gamma+1} = (0)$. Now if $c_\gamma = \xi_1 \circ \xi_2 \circ \ldots \xi_\gamma \neq 0$, then $\exp(c_\gamma) \neq 0$, since if

$$c_\gamma + \frac{c_\gamma^2}{2!} + \ldots + \frac{c_\gamma^{\nu-1}}{(\nu - 1)!} = 0, \tag{4.0.1}$$

and $c_\gamma^{\nu-1} \neq 0$, $c_\gamma^\nu = 0$, then, multiplying by $c_\gamma^{\nu-2}$, $c_\gamma^{\nu-1} = 0$, a contradiction. It follows, then, that if $\Lambda^\gamma \neq (0)$, $\mathfrak{G}_\gamma \neq (1)$. Certainly, however, $\mathfrak{G}_{\gamma+1} = (1)$ since c_γ, and therefore $c_\gamma + c_\gamma^2/2! + c_\gamma^3/3! + \ldots + c_\gamma^{\nu-1}/(\nu - 1)!$ is in the centre of Λ and so by (2.1) $\exp(c_\gamma)$ is in the centre of $\mathfrak{G}$. We have proved, therefore

THEOREM 4.1. *If A is a nil-algebra over a field of characteristic* 0, *then the class of $\Lambda(A)$ is equal to the class of $\mathfrak{G}(A)$, provided these classes are finite.*

REFERENCES

1. F. Hausdorff, *Die symbolische Exponentialformel in der Gruppentheorie*, Sitz. der Sächsischen Akad. Wiss. (Math.-phys. Klasse), *58* (1906), 19–48.

2. N. Jacobson, *The radical and semi-simplicity for arbitrary rings*, Amer. J. Math., *57* (1945), 300–320.

3. S. A. Jennings, *Rings whose associated Lie rings are nilpotent*, Bull. Am. Math. Soc., *53* (1947), 593–597.

TRANSACTIONS OF THE ROYAL SOCIETY OF CANADA
VOLUME XLIX : SERIES III : JUNE, 1955
SECTION THREE

On a Theorem of Poincaré and Malkin[1]

EUGENE LEIMANIS

Presented by R. L. JEFFERY, F.R.S.C.

1. A theorem of Poincaré and Malkin. The theorem in question is concerned with the existence of periodic solutions of a non-autonomous system of order n, containing a small parameter μ and being such that the generating system admits an infinity of periodic solutions depending upon k $(<n)$ arbitrary parameters. The case of a single parameter was considered by Poincaré (**1**, p. 84) and the general case with $1 < k$ $(<n)$ parameters by Malkin (**2**, p. 17). In this paper we shall consider an exceptional case of Malkin's theorem, when certain equations are satisfied identically.

Consider a system of differential equations

$$\frac{dx_i}{dt} = X_i(t; x_1, \ldots, x_n; \mu) \qquad (i = 1, 2, \ldots, n), \tag{1}$$

containing a small parameter μ. Assume that the X_i are analytic functions of their variables x_i in a certain domain D and of the small parameter μ at the point $\mu = 0$, but continuous periodic functions of the independent variable t with period T. The system (1) may then be written in the form

$$\frac{dx_i}{dt} = X_i^{(0)}(t; x_1, \ldots, x_n) + \mu X_i^{(1)}(t; x_1, \ldots, x_n) + \ldots, \tag{2}$$

where the $X_i^{(j)}$ $(i = 1, 2, \ldots, n;\ j = 0, 1, 2, \ldots)$ are analytic functions of the x_i in the domain D.

Assume that the generating system

$$\frac{dx_i}{dt} = X_i^{(0)}(t; x_1, \ldots, x_n) \qquad (i = 1, 2, \ldots, n), \tag{3}$$

obtained from (2) by putting $\mu = 0$, admits an infinity of periodic solutions

$$x_i^{(0)} = \phi_i(t; h_1, \ldots, h_k) \qquad (i = 1, 2, \ldots, n) \tag{4}$$

depending upon $1 \leqslant k$ $(<n)$ arbitrary parameters $h_1, \ldots, h_k$, and suppose that the generating solution under consideration corresponds to the values $h_j = h_j^*$ $(j = 1, 2, \ldots, k)$ of these parameters. Further assume that at least one of the determinants of order k of the matrix

$$\left\| \frac{\partial \phi_i}{\partial h_j} \right\|_{t=0,\, h_j = h_j^*} \qquad (i = 1, 2, \ldots, n;\ j = 1, 2, \ldots, k) \tag{5}$$

[1]This paper was prepared while the author attended the Summer Reseach Institute of the Canadian Mathematical Congress, Session 1955.

is different from zero. This means that the k initial conditions of the solution (4) can be taken arbitrarily, or, in other words, that the parameters $h_1, \ldots, h_k$ are independent.

For $\mu \neq 0$ we seek a periodic solution

$$x_i(t) = x_i(t; \beta_1, \ldots, \beta_n; \mu; \overset{*}{h}_1, \ldots, \overset{*}{h}_k) \tag{6}$$

of (1), satisfying the initial conditions

$$x_i(0) = \phi_i(0; \overset{*}{h}_1, \ldots, \overset{*}{h}_k) + \beta_i \tag{7}$$

and reducing for $\mu = 0$ to the generating solution

$$x_i^{(0)} = \phi_i(t; \overset{*}{h}_1, \ldots, \overset{*}{h}_k)$$

with the initial conditions

$$x_i^{(0)}(0) = \phi_i(0; \overset{*}{h}_1, \ldots, \overset{*}{h}_k).$$

The small parameters $\beta_i (i = 1, 2, \ldots, n)$ are to be determined in terms of μ in such a way that they vanish for $\mu = 0$ and the solution (6) becomes periodic in t with period T.

For this purpose it is necessary and sufficient that the following conditions be satisfied

$$\begin{aligned} \psi_i(\mu; \beta_1, \ldots, \beta_n; \overset{*}{h}_1, \ldots, \overset{*}{h}_k) &= x_i(T; \beta_1, \ldots, \beta_n; \mu; \overset{*}{h}_1, \ldots, \overset{*}{h}_k) \\ &\quad - x_i(0; \beta_1, \ldots, \beta_n; \mu; \overset{*}{h}_1, \ldots, \overset{*}{h}_k) \\ &= [x_i] = 0 \qquad (i = 1, 2, \ldots, n). \end{aligned} \tag{8}$$

From (4) and (7) it follows that for $\mu = 0$ the periodicity conditions (8) have a solution

$$\beta_i = \beta_i(h_1, \ldots, h_k) = \phi_i(0; h_1, \ldots, h_k) - \phi_i(0; \overset{*}{h}_1, \ldots, \overset{*}{h}_k) \tag{9}$$

which depends upon k arbitrary parameters $h_1, \ldots, h_k$. As a consequence of this, not only the Jacobian

$$\left| \frac{\partial \psi_i}{\partial \beta_j} \right|_{\mu = \beta_j = 0} \qquad (i, j = 1, 2, \ldots, n)$$

but all of its minors up to the order $n - k + 1$ inclusive, vanish. Suppose that at least one of the minors of order $n - k$ is different from zero. To be more specific, assume that

$$\left| \frac{\partial \psi_i}{\partial \beta_j} \right|_{\mu = \beta_j = 0} \neq 0, \qquad (i, j = 1, 2, \ldots, n - k). \tag{10}$$

Consequently, as Malkin (**2**, p. 18) has shown, among the determinants of order k of the matrix (5) which by assumption do not all vanish for $t = 0$, $h_j = h_j^*$, the determinant

$$\left| \frac{\partial \phi_r}{h \partial_s} \right|_{t=0, h_s = \overset{*}{h}_s} \neq 0 \qquad (r = n - k + 1, \ldots, n; s = 1, \ldots, k). \tag{11}$$

Since the determinant (10) is different from zero, the first $n - k$ equations of (8) can be solved uniquely for $\beta_1, \ldots, \beta_{n-k}$ as power series in $\beta_{n-k+1}, \ldots, \beta_n$ and μ, vanishing for $\beta_{n-k+1} = \ldots = \beta_n = \mu = 0$. Suppose that these solutions are substituted in the last k equations of the system (8). These latter then will become equations in $\mu; \beta_{n-k+1}, \ldots, \beta_n; h_1^*, \ldots, h_k^*$ which may be written in the form

$$F_j(\mu; \beta_{n-k+1}, \ldots, \beta_n; h_1^*, \ldots, h_k^*) = 0 \qquad (j = 1, 2, \ldots, k), \tag{12}$$

where the F_j are analytic functions of $\mu; \beta_{n-k+1}, \ldots, \beta_n$, and they vanish with these quantities.

Equations (8) admit for $\mu = 0$ a solution (9) which depends on the k arbitrary parameters $h_1, \ldots, h_k$. Moreover, since the determinant (11) is different from zero, the quantities $\beta_{n-k+1}, \ldots, \beta_n$ themselves can be taken as those parameters. Consequently, for $\mu = 0$ equations (12) must be satisfied for arbitrary values of the parameters $\beta_{n-k+1}, \ldots, \beta_n$. Therefore the F_j must have the form

$$F_j = \mu\, \Phi_j(\mu; \beta_{n-k+1}, \ldots, \beta_n; h_1^*, \ldots, h_k^*), \tag{12'}$$

where the Φ_j are analytic functions in $\mu, \beta_{n-k+1}, \ldots, \beta_n$, but do not vanish, in general, for $\mu = \beta_{n-k+1} = \ldots = \beta_n = 0$. Hence, after the factor μ is divided out, equations (12) become

$$\Phi_j = P_j + Q_{j1}\beta_{n-k+1} + \ldots + Q_{jk}\beta_n + R_j\mu + \ldots = 0, \tag{13}$$

where the coefficients $P_j, Q_{js}, R_j, \ldots$ $(s = 1, 2, \ldots, k)$ depend on the values $h_1^*, \ldots, h_k^*$ of the parameters entering in the generating solution. In order that equations (13) have a solution for $\beta_{n-k+1}, \ldots, \beta_n$ which vanishes with μ, it is necessary that

$$P_j(h_1^*, \ldots, h_k^*) = 0 \qquad (j = 1, 2, \ldots, k). \tag{14}$$

If, in addition, the determinant

$$|Q_{js}(h_1^*, \ldots, h_k^*)| \neq 0 \qquad (j, s = 1, 2, \ldots, k), \tag{15}$$

the system (13) can be solved uniquely for $\beta_{n-k+1}, \ldots, \beta_n$ in terms of μ. This solution is analytic in μ and vanishes with μ. It can be shown **(2**, p. 22**)** that the condition (15) is equivalent to the condition

$$\left| \frac{\partial P_j}{\partial h_s^*} \right| \neq 0 \qquad (j, s = 1, 2, \ldots, k), \tag{16}$$

i.e., the solution $h_1^*, \ldots, h_k^*$ is a simple root of the system (14). These are the essentials of the theorem of Malkin.

In this paper the case when the equations (14) are satisfied identically will be considered. This latter case is of importance in some applications to the theory of nonlinear oscillations.

2. An extension of the previous theorem. The following way of constructing periodic solutions of a non-autonomous system (1) is suggested by the method of proving their existence. It is necessary to write down, first, the functions (6); second, the periodicity conditions (8); third, to solve these equations for β_i $(i = 1, 2, \ldots, n)$ in terms of μ, by showing that the solutions so obtained are power series in μ, say,

$$\beta_i = \sum_{\alpha=1}^{\infty} \beta_i^{(\alpha)} \mu^\alpha;$$

and, finally, to substitute these solutions in (6) and to show that the coefficients $\beta_i^{(\alpha)}$ $(i = 1, 2, \ldots, n; \alpha = 1, 2, \ldots)$ of the power series for β_i can be uniquely determined in such a way that the $x_i(t)$ given by (6) become periodic in t with period T. These four steps will be carried out in the following pages.

(i) According to a theorem of Poincaré (**1**, p. 58) the $x_i(t)$ given by (6) are expansible in converging power series of μ; $\beta_1, \ldots, \beta_n$ when the absolute values of these $n+1$ quantities are sufficiently small. The coefficients are continuous periodic functions of t with period T. Therefore let us look for the $x_i(t)$ having the form

$$x_i = \sum_{\alpha=0}^{\infty} x_{i\alpha}\mu^\alpha \qquad (i = 1, 2, \ldots, n), \tag{17}$$

where the coefficients $x_{i\alpha}$ of the various powers of μ are all expansible as converging power series in $\beta_1, \ldots, \beta_n$, i.e.,

$$x_{i\alpha} = x_i^{(\alpha)} + \sum_{\nu=1}^{\infty} \sum_{i_1, \ldots, i_\nu = 1}^{n} x_{i i_1 \ldots i_\nu}^{(\alpha)} \beta_{i_1} \ldots \beta_{i_\nu} \tag{18}$$

$$(i = 1, \ldots, n; \alpha = 0, 1, \ldots).$$

Substituting (17) in (2) and equating the coefficients of corresponding powers of μ in the two members it is found that

$$\frac{\partial x_{i0}}{\partial t} = X_i^{(0)}(t; x_{10}, \ldots, x_{n0}) \qquad (i = 1, \ldots, n),$$

$$\frac{\partial x_{i1}}{\partial t} = X_i^{(1)}(t; x_{10}, \ldots, x_{n0}) + \sum_{j=1}^{n} \left(\frac{\partial X_i^{(0)}}{\partial x_j}\right)_0 x_{j1},$$

$$\frac{\partial x_{i2}}{\partial t} = \tfrac{1}{2} X_i^{(*)}(t; x_{10}, \ldots, x_{n0}) + \sum_{j=1}^{n} \left(\frac{\partial X_i^{(1)}}{\partial x_j}\right)_0 x_{j1},$$

$$+ \tfrac{1}{2} \sum_{j,k=1}^{n} \left(\frac{\partial^2 X_i^{(0)}}{\partial x_j \partial x_k}\right)_0 x_{j1}x_{k1} + \sum_{j=1}^{n} \left(\frac{\partial X_i^{(0)}}{\partial x_j}\right)_0 x_{j2},$$

. . . .

$$\frac{dx_{i\alpha}}{dt} = X_{i\alpha} + \sum_{j=1}^{n} \left(\frac{\partial x_i^{(0)}}{\partial x_j}\right)_0 x_{j\alpha},$$

where $X_{i\alpha}$ is a polynomial in $x_{j0}, \ldots, x_{j,\alpha-1}$ $(j = 1, \ldots, n)$ with periodic coefficients of period T. After the x_{i0} have been determined, the remaining $x_{i\alpha}$ $(\alpha = 1, 2, \ldots)$ depend upon the solution of systems of linear nonhomogeneous differential equations. Assuming that the $x_{j0}, \ldots, x_{j,\alpha-1}$ have been determined in the previous steps, the $X_{i\alpha}$ are known periodic functions of t with period T. Parantheses with the subscript 0 denote that after differentiation the variables x_j have been replaced by x_{j0}.

Substituting the series (17) and (18) into the differential equations (2) and equating coefficients of like powers of $\beta_1, \ldots, \beta_n, \mu$, it is found that the $x_i^{(\alpha)}$ satisfy the same differential equations as the $x_{i\alpha}$ $(\alpha = 0, 1, \ldots)$; in particular, the $x_i^{(0)}$ satisfy the generating system (3). The initial conditions are

(19) $$x_i^{(0)}(0) = \phi_i(0; \overset{*}{h}_1, \ldots, \overset{*}{h}_k),\ x_i^{(\alpha)}(0) = 0 \qquad (\alpha = 1, 2, \ldots).$$

The $x_{ii_1}^{(0)}$ satisfy the variational equations of the generating system

$$\frac{\partial x_{ii_1}^{(0)}}{\partial t} = \sum_{j=1}^{n} \left[\frac{\partial X_i^{(0)}(t; x_1, \ldots, x_n)}{\partial x_j} \right]_0 x_{ji_1}^{(0)}$$

with the initial conditions

(20) $$x_{ii_1}^{(0)}(0) = \begin{cases} 0 \text{ if } i_1 \neq i \\ 1 \text{ if } i_1 = i. \end{cases}$$

Finally the $x_{ii_1}^{(\alpha)}{}_{\ldots i_\nu}$ satisfy linear nonhomogeneous differential equations with periodic coefficients

$$\frac{dx_{ii_1\ldots i_\nu}^{(\alpha)}}{dt} = X_{ii_1\ldots i_\nu}^{(\alpha)} + \sum_{j=1}^{n} \left[\frac{\partial X_i^{(0)}(t; x_1, \ldots, x_n)}{\partial x_j} \right]_0 x_{ji_1\ldots i_\nu}^{(\alpha)},$$

the initial conditions being

(21) $$x_{ii_1\ldots i_\nu}^{(\alpha)}(0) = 0$$

$(\nu = 1, 2, \ldots, n; \alpha = 0, 1, 2, \ldots$; except the case $\alpha = 0, \nu = 1$, given by (20)). Square brackets with the subscript 0 denote that after differentiation the variables x_j have been replaced by $x_j^{(0)}$.

The $X^{(\alpha)}{}_{ii_1\ldots i_\nu}$ are polynomials of those $x^{(\gamma)}{}_{ii_1\ldots i_\lambda}$ for which $\gamma+\lambda<\alpha+\nu$. Assuming that these last functions have been calculated in the previous steps, the $X^{(\alpha)}{}_{ii_1\ldots i_\nu}$ are known periodic functions of t with period T.

(ii) The last k conditions of periodicity (8) according to (12) and (12′) can be written in the form

(22) $$\Phi_j(\mu; \beta_{n-k+1}, \ldots, \beta_n; \overset{*}{h}_1, \ldots, \overset{*}{h}_k) = 0 \qquad (j = 1, \ldots, k),$$

where

(23) $$\Phi_j = \sum_{\alpha=0}^{\infty} \Phi_{j\alpha}\mu^\alpha,$$

$$\Phi_{j0} = \Phi_j^{(0)} + \sum_{\nu=1}^{\infty} \sum_{i_1, \ldots, i_\nu=n-k+1}^{n} \Phi_{ji_1\ldots i_\nu}^{(0)} \beta_{i_1} \cdots \beta_{i_\nu},$$

$$\Phi_{j\alpha} = \Phi_j^{(\alpha)} + \sum_{\nu=1}^{\infty} \sum_{i_1, \ldots, i_\nu=n-k+1}^{n} \Phi_{ji_1\ldots i_\nu}^{(\alpha)} \beta_{i_1} \cdots \beta_{i_\nu}.$$

The $x_{i\alpha}$ and $x_i^{(\alpha)}$ satisfy the same differential equations, as we have seen before. Since the determinant (11) is different from zero, introduce in $\Phi_{j\alpha}$ by means of the last k relations (9) the parameters $h_1, \ldots, h_k$ instead of $\beta_{n-k+1}, \ldots, \beta_n$, and denote the new function by $\bar{\Phi}_{j\alpha}(h_1, \ldots, h_k; h_1^*, \ldots, h_k^*)$. Then, because $h_j = h_j^*$ correspond to $\beta_{n-k+1} = \ldots = \beta_n = 0$, we have

$$(24) \qquad \bar{\Phi}_{j\alpha}(h_1, \ldots, h_k; \overset{*}{h_1}, \ldots, \overset{*}{h_k})$$

$$= \Phi_{j\alpha}[\beta_{n-k+1}(h_1, \ldots, h_k), \ldots, \beta_n(h_1, \ldots, h_k); \overset{*}{h_1}, \ldots, \overset{*}{h_k}] = \Phi_j^{(\alpha)}(h_1, \ldots, h_k).$$

From (12′) it follows that

$$\Phi_{j0} = 0,$$

i.e., the equations

$$\Phi_j^{(0)} = \Phi_{ji_1\ldots i_\nu}^{(0)} = 0$$

are satisfied identically with respect to $h_1^*, \ldots, h_k^*$. On the other hand comparison of (23) and (13) shows that the following relations hold

$$\begin{aligned} \Phi_j^{(1)} &= P_j, \\ \Phi_{j,n-k+s}^{(1)} &= Q_{js}, \qquad (j, s = 1, 2, \ldots, k). \\ \Phi_j^{(2)} &= R_j. \end{aligned}$$

According to the purpose of this paper assume now that the equations

$$(25) \qquad \Phi_j^{(\gamma)} = 0 \qquad (\gamma = 1, 2, \ldots, m-1)$$

are identically satisfied with respect to $h_1^*, \ldots, h_k^*$, but the system of equations

$$(26) \qquad \Phi_j^{(m)} = 0 \qquad (j = 1, 2, \ldots, k)$$

has a solution $h_1^*, \ldots, h_k^*$ for which the Jacobian

$$(27) \qquad \left| \frac{\partial \Phi_j^{(m)}}{\partial \overset{*}{h_s}} \right| \neq 0, \qquad (j, s = 1, 2, \ldots, k).$$

(iii) It remains now to show that under the assumptions made above the periodicity conditions (22) can be solved for the β_i ($i = n - k + 1, \ldots, n$) in power series of μ, convergent for $|\mu|$ sufficiently small and vanishing with μ.

From the assumption (25) it follows that the equations

$$\Phi_{j\gamma}(\beta_{n-k+1}, \ldots, \beta_n; \overset{*}{h_1}, \ldots, \overset{*}{h_k}) = 0$$

are identically satisfied with respect to their variables; or, that the equations

$$\Phi_{ji_1\ldots i_\nu}^{(\gamma)}(\overset{*}{h_1}, \ldots, \overset{*}{h_k}) = 0$$

are identically satisfied in $h_1^*, \ldots, h_k^*$. Proof by mathematical induction runs as follows.

Assume that the equations

$$(28)\qquad \Phi_{j\gamma}(\beta_{n-k+1}, \ldots, \beta_n; \overset{*}{h_1}, \ldots, \overset{*}{h_k}) = 0 \qquad (\gamma = 1, 2, \ldots, r-1; r < m)$$

are satisfied, and let us show that under the supplemental assumption

$$(29)\qquad \Phi_j^{(r)}(\overset{*}{h_1}, \ldots, \overset{*}{h_k}) = 0$$

we have

$$(30)\qquad \Phi_{jr}(\beta_{n-k+1}, \ldots, \beta_n; \overset{*}{h_1}, \ldots, \overset{*}{h_k}) = 0.$$

In fact, the function Φ_j, given by (23), because of (28) becomes

$$(31)\qquad \Phi_j = \sum_{\alpha=r}^{\infty} \Phi_{j\alpha}\mu^{\alpha}.$$

Substitute (31) into (22) and divide out the factor μ^r. The conditions of periodicity then read

$$(32)\qquad \Psi_{jr}(\mu; \beta_{n-k+1}, \ldots, \beta_n; \overset{*}{h_1}, \ldots, \overset{*}{h_k}) = \Phi_{jr} + \sum_{\alpha=1}^{\infty} \Phi_{j,r+\alpha}\mu^{\alpha} = 0$$
$$(j = 1, 2, \ldots, k).$$

Introduce in Ψ_{jr} by means of the last k relations (9) the parameters $h_1, \ldots, h_k$ instead of $\beta_{n-k+1}, \ldots, \beta_n$, and denote the new function by $\bar{\Psi}_{jr}(\mu; h_1, \ldots, h_k; h_1^*, \ldots, h_k^*)$. Then

$$\bar{\Psi}_{jr}(0; h_1, \ldots, h_k; \overset{*}{h_1}, \ldots, \overset{*}{h_k}) = \bar{\Phi}_{jr}(h_1, \ldots, h_k; \overset{*}{h_1}, \ldots, \overset{*}{h_k})$$

and because of (24) and (29) we have

$$\bar{\Phi}_{jr}(h_1, \ldots, h_k; \overset{*}{h_1}, \ldots, \overset{*}{h_k}) = 0,$$

i.e., (30) holds. By our hypothesis (25), the result holds for $r = 1$, and therefore it holds for all r.

Let $r = m - 1$. Then, after dividing out in (31) the factor μ^m, the periodicity conditions assume the form

$$(33)\qquad \Phi_{jm} + \sum_{\alpha=1}^{\infty} \Phi_{j,m+\alpha}\mu^{\alpha} = 0,$$

where

$$(34)\qquad \Phi_{jm} = \Phi_j^{(m)}(\overset{*}{h_1}, \ldots, \overset{*}{h_k}) + \sum_{i_1=n-k+1}^{n} \Phi_{ji_1}^{(m)}(\overset{*}{h_1}, \ldots, \overset{*}{h_k})\beta_{i_1}$$
$$+ \sum_{\nu=2}^{\infty} \sum_{i_1,\ldots,i_\nu=1}^{n} \Phi_{ji_1\ldots i_\nu}^{(m)}(\overset{*}{h_1}, \ldots, \overset{*}{h_k})\, \beta_{i_1} \ldots \beta_{i_\nu}.$$

In order that the system (33) should ensure the existence of the desired solution for the β_i $(i = n - k + 1, \ldots, n)$ in terms of μ, it is necessary that $h_1^*, \ldots, h_k^*$ satisfy the equations (26). If, in addition, the determinant

$$(35)\qquad |\Phi_{ji_1}^{(m)}(\overset{*}{h_1}, \ldots, \overset{*}{h_k})| \neq 0,$$

equations (33) will possess a unique analytic solution for the β_i $(i = n - k + 1, \ldots, n)$ vanishing with μ.

To prove (35) let us introduce once more the parameters $h_1, \ldots, h_k$ instead of $\beta_{n-k+1}, \ldots, \beta_n$ in the function $\Phi_{jm}(\beta_{n-k+1}, \ldots, \beta_n; h_1^*, \ldots, h_k^*)$ and denote the new function by $\bar{\Phi}_{jm}(h_1, \ldots, h_k; h_1^*, \ldots, h_k^*)$. Then by (24) we have

$$\bar{\Phi}_{jm}(h_1, \ldots, h_k; h_1^*, \ldots, h_k^*) = \phi_j^{(m)}(h_1, \ldots, h_k)$$

and hence

$$\frac{\partial \bar{\Phi}_{jm}}{\partial h_s^*} = 0 \qquad (s = 1, 2, \ldots, k). \tag{36}$$

Remembering that $h_s = h_s^*$ corresponds to $\beta_{n-k+1} = \ldots = \beta_n = 0$, it follows from (9), (34) and (36) that

$$\frac{\partial \Phi_j^{(m)}}{\partial h_s^*} - \sum_{i_1=n-k+1}^{n} \Phi_{ji_1}^{(m)} \frac{\partial \Phi_{i_1}}{\partial h_s^*} = 0 \qquad (j, s = 1, 2, \ldots, k). \tag{37}$$

From (37) it is easily seen that

$$\left| \frac{\partial \Phi_j^{(m)}}{\partial h_s^*} \right| = \left| \Phi_{ji_1}^{(m)} \right| \left| \frac{\partial \Phi_{i_1}}{\partial h_s^*} \right| \qquad (i_1 = n - k + 1, \ldots, n)$$

and consequently

$$\left| \Phi_{ji_1}^{(m)} \right| = \left| \frac{\partial \Phi_j^{(m)}}{\partial h_s^*} \right| : \left| \frac{\partial \Phi_{i_1}}{\partial h_s^*} \right| .$$

Because of (11) the denominator in the last formula is different from zero, and hence the desired result (35) is established.

Consequently the periodicity conditions (22) can be solved uniquely for the β_i in the form

$$\beta_i = \sum_{\alpha=1}^{\infty} \beta_i^{(\alpha)} \mu^\alpha, \qquad (i = n - k + 1, \ldots, n) \tag{38}$$

where

$$\beta_i^{(\alpha)} = \beta_i^{(\alpha)}(h_1^*, \ldots, h_k^*).$$

Substituting (38) in the formulas obtained by solving the first $n - k$ equation (8) for β_i $(i = 1, 2, \ldots, n - k)$ in terms of $\beta_{n-k+1}, \ldots, \beta_n, \mu$ we obtain

$$\beta_i = \sum_{\alpha=1}^{\infty} \beta_i^{(\alpha)} \mu^\alpha, \qquad (i = 1, 2, \ldots, n - k) \tag{39}$$

where

$$\beta_i^{(\alpha)} = \beta_i^{(\alpha)}(h_1^*, \ldots, h_k^*).$$

(iv) Substitute (39) and (38) in (17) and (18) and denote the solution of (1) so obtained by

$$\bar{x}_i = \sum_{\alpha=0}^{\infty} \bar{x}_{i\alpha} \mu^\alpha, \qquad (i = 1, 2, \ldots, n). \tag{40}$$

The expressions of the coefficients $\bar{x}_{i\alpha}$ in terms of the $x^{(\alpha)}{}_{i i_1 \ldots i_\nu}$ and the $\beta_i{}^{(\alpha)}$ are as follows

$$\bar{x}_{i0} = x_i^{(0)},$$

$$\bar{x}_{i1} = x_i^{(1)} + \sum_{i_1=1}^{n} x_{i i_1}^{(0)} \beta_{i_1}^{(1)},$$

$$\bar{x}_{i\alpha} = x_i^{(\alpha)} + \sum_{s=1}^{\alpha-1} \sum_{\nu=1}^{s} \sum_{i_1,\ldots,i_\nu=1}^{n} x_{i i_1 \ldots i_\nu}^{(\alpha-s)} \sum_{\alpha_1+\ldots+\alpha_\nu=s} \beta_{i_1}^{(\alpha_1)} \ldots \beta_{i_\nu}^{(\alpha_\nu)}$$

$$+ \sum_{i_1=1}^{n} x_{i i_1}^{(0)} \beta_{i_1}^{(\alpha)} + \sum_{\nu=2}^{\alpha} \sum_{i_1,\ldots,i_\nu=1}^{n} x_{i i_1 \ldots i_\nu}^{(0)} \sum_{\alpha_1+\ldots+\alpha_\nu=\alpha} \beta_{i_1}^{(\alpha_1)} \ldots \beta_{i_\nu}^{(\alpha_\nu)}$$

$$(i = 1, 2, \ldots, n; \alpha = 2, 3, \ldots).$$

From (19), (20) and (21) we obtain the following initial conditions for the $\bar{x}_{i\alpha}$:

$$\text{(41)} \qquad \begin{aligned} \bar{x}_{i0}(0) &= x_i^{(0)}(0) = \phi_i(0; h_1^*, \ldots, h_k^*), \\ \bar{x}_{i\alpha}(0) &= \beta_i^{(\alpha)}, \qquad (i = 1, 2, \ldots, n; \alpha = 1, 2, \ldots). \end{aligned}$$

Taking the solution of (1) in the form of the power series (40), the coefficients $\bar{x}_{i\alpha}$ will satisfy the same linear nonhomogeneous differential equations as the coefficients $x_i^{(\alpha)}$, except that the initial conditions (19) are replaced by (41). The problem is to show that the $\beta_i^{(\alpha)}$ can be uniquely determined in such a way that the solution (40) becomes periodic in t with period T. The periodicity conditions of the solution (40) are

$$[\bar{x}_i] = 0 \qquad (i = 1, 2, \ldots, n),$$

or, equivalently,

$$\text{(42)} \qquad [\bar{x}_{i\alpha}] = 0, \quad (i = 1, 2, \ldots, n; \alpha = 0, 1, 2, \ldots).$$

The set of conditions obtained from (42) for $\alpha = 0$ is identically satisfied with respect to $h_1, \ldots, h_k$. The remaining conditions may be written as

$$\text{(43)} \qquad [\bar{x}_{i1}] = \psi_i^{(1)} + \sum_{i_1=1}^{n-k} \psi_{i i_1}^{(0)} \beta_{i_1}^{(1)} + \sum_{i_1=n-k+1}^{n} \psi_{i i_1}^{(0)} \beta_{i_1}^{(1)} = 0,$$

$$\text{(44)} \qquad [\bar{x}_{i\alpha}] = \psi_i^{(\alpha)} + \sum_{s=1}^{\alpha-1} \sum_{\nu=1}^{s} \sum_{i_1,\ldots,i_\nu=1}^{n} \psi_{i i_1 \ldots i_\nu}^{(\alpha-s)} \sum_{\alpha_1+\ldots+\alpha_\nu=s} \beta_{i_1}^{(\alpha_1)} \ldots \beta_{i_\nu}^{(\alpha_\nu)}$$

$$+ \sum_{i_1=1}^{n-k} \psi_{i i_1}^{(0)} \beta_{i_1}^{(\alpha)} + \sum_{i_1=n-k+1}^{n} \psi_{i i_1}^{(0)} \beta_{i_1}^{(\alpha)}$$

$$+ \sum_{\nu=2}^{\alpha} \sum_{i_1,\ldots,i_\nu=1}^{n} \psi_{i i_1 \ldots i_\nu}^{(0)} \sum_{\alpha_1+\ldots+\alpha_\nu=\alpha} \beta_{i_1}^{(\alpha_1)} \ldots \beta_{i_\nu}^{(\alpha_\nu)} = 0,$$

$$(i = 1, 2, \ldots, n; \alpha = 2, 3, \ldots).$$

The sets of the first $n - k$ equations (43) and (44) are linear in

$$\beta_{i_1}^{(1)} \text{ and } \beta_{i_1}^{(\alpha)} \ (i_1 = 1, 2, \ldots, n - k).$$

Therefore, because of the assumption (10) which can be written in the form

$$|\psi_{ii_1}^{(0)}| \neq 0 \qquad (i, i_1 = 1, 2, \ldots, n-k),$$

they can be solved uniquely with respect to those variables, i.e. with respect to the initial values $\bar{x}_{i\alpha}(0)$ $(i = 1, 2, \ldots, n-k; \alpha = 1, 2, \ldots)$ as functions of $h_1, \ldots, h_k$ and $\bar{x}_{n-k+1,\alpha}(0), \ldots, \bar{x}_{n\alpha}(0)$.

Substituting these expressions in the remaining sets of k conditions (43) and (44) we obtain

$$(45) \qquad [\bar{x}_{j\gamma}] = \Phi_j^{(\gamma)} = 0, \quad (\gamma = 1, 2, \ldots, m-1; j = 1, 2, \ldots, k)$$

$$(46) \qquad [\bar{x}_{jm}] = \Phi_j^{(m)} = 0,$$

$$(47) \qquad [\bar{x}_{j,m+1}] = \Phi_j^{(m+1)} + \sum_{i_1=n-k+1}^{n} \Phi_{ji_1}^{(m)} \beta_{i_1}^{(1)} = 0,$$

$$(48) \qquad [\bar{x}_{j,m+\alpha}] = \Phi_j^{(m+\alpha)} + \sum_{s=1}^{\alpha-1} \sum_{\nu=1}^{s} \sum_{i_1,\ldots,i_\nu=1}^{n} \Phi_{ji_1\ldots i_\nu}^{(m+\alpha-s)} \sum_{\alpha_1+\ldots+\alpha_\nu=s} \beta_{i_1}^{(\alpha_1)} \ldots \beta_{i_\nu}^{(\alpha_\nu)}$$
$$+ \sum_{i_1=n-k+1}^{n} \Phi_{ji_1}^{(m)} \beta_{i_1}^{(\alpha)}$$
$$+ \sum_{\nu=2}^{\alpha} \sum_{i_1,\ldots,i_\nu=n-k+1}^{n} \Phi_{ji_1\ldots i_\nu}^{(m)} \sum_{\alpha_1+\ldots+\alpha_\nu=\alpha} \beta_{i_1}^{(\alpha_1)} \ldots \beta_{i_\nu}^{(\alpha_\nu)} = 0$$
$$(\alpha = 2, 3, \ldots).$$

The equations (45) according to our assumption (25) are identically satisfied. In view of (27), the equations (46) determine the values $h_1^*, \ldots, h_k^*$ of the parameters $h_1, \ldots, h_k$. Finally the equations (47) and (48) are linear in $\beta_{i_1}^{(1)}$, $\beta_{i_1}^{(\alpha)}$ $(i_1 = n-k+1, \ldots, n)$, and because of (35) they can be solved with respect to these variables, i.e. with respect to the initial values $\bar{x}_{j\alpha}(0)$ $(j = n-k+1, \ldots, n;\ \alpha = 2, 3, \ldots)$. This completes the construction of the desired periodic solution of (1).

The results obtained can be formulated as the following extended theorem of Poincaré and Malkin.

THEOREM. *Given the system* (1), *assume that the generating system* (3) *admits an infinity of periodic solutions* (4) *depending upon k arbitrary parameters* $h_1, \ldots, h_k$. *Further assume that the required periodic solution of the given system* (1) *can be represented as a power series* (40) *in* μ *which for* $\mu = 0$ *reduces to the generating solution* (4). *Moreover, assume that the equations* (45) *are satisfied identically with respect to the parameters, and that the equations* (46) *have a simple solution* $h_j = h_j^*(j = 1, 2, \ldots, k)$. *Under these assumptions there exists a uniquely determined periodic solution which is analytic in* μ *at the point* $\mu = 0$.

REFERENCES

1. H. Poincaré, *Les méthodes nouvelles de la Mécanique Céleste*, vol. 1 (Paris, Gauthier-Villars, 1892).
2. I. G. Malkin, *The methods of Lyapunov and Poincaré in the theory of nonlinear oscillations*, Gos. Izdat. Teh.-Teor. Lit., Leningrad-Moscow, 1949 (Russian).

TRANSACTIONS OF THE ROYAL SOCIETY OF CANADA
VOLUME XLIX : SERIES III : JUNE, 1955
SECTION THREE

An Asymptotic Formula for the Bell Numbers

LEO MOSER and MAX WYMAN, F.R.S.C.

1. Introduction. Properties of the Bell numbers G_n, defined by

$$\sum_{n=0}^{\infty} \frac{G_n z^n}{n!} = e^{e_z - 1} \tag{1.1}$$

have been studied by many authors. A recent thesis of Finlayson **(5)** lists over fifty references to these numbers. The problem of determining a formula for the asymptotic behaviour of G_n for large n has been suggested several times. However, as far as we are aware, only two such formulae have been derived. Knopp **(6)** gives the formula

$$\frac{G_n}{n!} = \left(\frac{1+\eta_n}{\log n}\right)^n \tag{1.2}$$

where $\eta_n \to 0$ as $n \to \infty$. Since the way in which η_n approaches zero is not specified the formula is of no value for computation and does not constitute an asymptotic expansion for G_n. Epstein **(4)** gives the formula

$$G_n \sim \left(\frac{n\, e^{1/\log n}}{\log n}\right)^n \tag{1.3}$$

As we shall see, this result is in error. The method used to obtain it is rather long, but could be used to get a correct expression for the first term of an asymptotic expansion for G_n. In the present paper a complete asymptotic expansion for the Bell numbers will be obtained by an entirely different method.

2. Asymptotic expansion. Since the iterated exponential function occurs throughout the paper we shall use both e^z and $\exp z$ to denote the exponential function.

From (1.1) and Cauchy's theorem

$$G_n = \frac{n!}{2\pi i} \int_c [\exp(e^z - 1)]/z^{n+1} \cdot dz, \tag{2.1}$$

where C is the circle $z = Re^{i\theta}$. Hence,

$$G_n = \frac{n!}{2\pi e R^n} \int_{-\pi}^{\pi} \exp[\exp(Re^{i\theta}) - in\theta]\, d\theta. \tag{2.2}$$

This can be put in the form

$$G_n = A \int_{-\pi}^{\pi} \exp(F(\theta))\, d\theta, \tag{2.3}$$

where

$$A = [n!\exp(e^R - 1)]/2\pi R^n \tag{2.4}$$

and

$$F(\theta) = \exp(Re^{i\theta}) - in\theta - \exp R. \tag{2.5}$$

Let us define ϵ by

$$\epsilon = e^{-3R/8} \tag{2.6}$$

and consider the integral J, defined by

$$J = \int_{\epsilon}^{\pi} \exp(F(\theta))\, d\theta. \tag{2.7}$$

An easy computation proves the existence of a constant $k > 0$ such that

$$|J| < \exp(-k\, Re^{R/4}) \tag{2.8}$$

Since we shall show that the terms of this order may be neglected in our asymptotic expansion, we shall anticipate this result and use (2.3) to obtain

$$G_n \sim A \int_{-\epsilon}^{\epsilon} \exp(F(\theta))\, d\theta. \tag{2.9}$$

Our next step is to expand $F(\theta)$, as given by (2.5) in a Maclaurin expansion about $\theta = 0$. If we introduce the operator Θ defined by

$$\Theta = R\frac{d}{dR} \tag{2.10}$$

we may write

$$F(\theta) = (Re^R - n)\, i\theta - \tfrac{1}{2}(R^2 + R)\, e^R\theta^2 + \sum_{k=3}^{\infty} (\Theta^k e^R)\frac{(i\theta)^k}{k!}. \tag{2.11}$$

At this stage we choose R to be the unique real solution of the equation

$$Re^R = n. \tag{2.12}$$

With this choice (2.11) becomes

$$F(\theta) = -\tfrac{1}{2}(R^2 + R)\, e^R\theta^2 + \sum_{k=3}^{\infty} (\Theta^k e^R)\frac{(i\theta)^k}{k!}. \tag{2.13}$$

We now introduce the following notation:

$$\phi = [\tfrac{1}{2}(R^2 + R)\, e^R]^{\frac{1}{2}}\theta \tag{2.14}$$

$$B = A/[\tfrac{1}{2}(R^2 + R)\, e^R]^{\frac{1}{2}} \tag{2.15}$$

$$h = \epsilon[\tfrac{1}{2}(R^2 + R)\, e^R]^{\frac{1}{2}} = \tfrac{1}{2}(R^2 + R)^{\frac{1}{2}}e^{R/8} \tag{2.16}$$

$$a_k = e^{-R}\, \Theta^{k+2}(e^R)(i\phi)^{k+2}/(k + 2)![\tfrac{1}{2}(R^2 + R)]^{k+2/2} \tag{2.17}$$

$$z = e^{-R/2} \tag{2.18}$$

$$f(z) = \sum_{k=1}^{\infty} a_k z^k. \tag{2.19}$$

We note in passing that the a_k are polynomials in ϕ.

Making the substitution (2.14) in (2.9) we find

$$G_n \sim B\int_{-h}^{h} e^{-\phi^2+f(z)}d\phi. \tag{2.20}$$

From (2.17) an easy calculation shows that there exists a fixed R_0 such that for all $R > R_0$,

$$|a_k| < |2\phi|^{k+2}. \tag{2.21}$$

We have defined z as a function of R. However for the moment we consider z to be an independent variable and expand $e^{f(z)}$ in a convergent Maclaurin expansion of the form

$$e^{f(z)} = \sum_{k=0}^{\infty} b_k z^k, \quad b_0 = 1. \tag{2.22}$$

Further, by (2.21), $z = e^{-R/2}$ is inside the domain of convergence. Hence at this stage we may again take $z = e^{-R/2}$. We note in passing that the b_k are polynomials in ϕ. Further, b_{2k+1} contains only odd powers of ϕ and b_{2k} only even powers of ϕ.

Using (2.22) we may write (2.20) in the form

$$G_n = B\left\{\sum_{k=0}^{s-1}\left(\int_{-h}^{h} e^{-\phi^2}b_k\,d\phi\right)z^k + R_s\right\} \tag{2.24}$$

where

$$R_s = \int_{-h}^{h} e^{-\phi^2}\left(\sum_{k=s}^{\infty} b_k z^k\right)d\phi. \tag{2.25}$$

From (2.12) we see that $R \to \infty$ as $n \to \infty$. Further, (2.16) implies that $h \to \infty$ as $R \to \infty$. From these facts and the known asymptotic expansion of functions of the form $\int_{-h}^{h} e^{-\phi^2}$. (polynomial in ϕ) $d\phi$, the replacement of h by ∞ in (2.24) is easily justified. Hence

$$G_n \sim B\left\{\sum_{k=0}^{s-1}\left(\int_{-\infty}^{\infty} e^{-\phi^2}b_k d\phi\right)z^k + R_s\right\}. \tag{2.25}$$

In order to complete our proof we must show that $R_s = 0(|z|^s)$. By a lemma of the authors **(7)**, (2.21) implies

$$|b_k| \leqslant |2\phi|^{k+2}(1 + |2\phi|^2)^{k-1}. \tag{2.25}$$

This yields

$$\left|\sum_{k=s}^{\infty} b_k z^k\right| \leqslant P_s(|\phi|)\,|z|^s/M, \tag{2.27}$$

where $P_s(|\phi|)$ is a polynomial in $|\phi|$ and M is given by

$$M = 1 - |z|\,|2\phi|\,(1 + 2|\phi|^2). \tag{2.28}$$

Since $|\phi| < h$ and $z = e^{-R/2}$, $|\phi|^3|z| < [\frac{1}{2}(R^2 + R)]^{3/2}e^{-R/8} \to 0$ as $R \to \infty$. Hence for large R, $M > \frac{1}{2}$. Finally,

$$\int_{-\infty}^{\infty} e^{-\phi^2} P_s(|\phi|)\, d\phi$$

exists, which implies $|R_s| = O(|z|^s)$. This completes the proof that

$$G_n \sim B \sum_{k=0}^{\infty} \left(\int_{-\infty}^{\infty} e^{-\phi^2} b_k d\phi \right) e^{-kR/2}. \tag{2.29}$$

We have noted that the b_{2k+1}, as polynomials in ϕ, contain only odd powers. Hence

$$G_n \sim B \sum_{k=0}^{\infty} \left(\int_{-\infty}^{\infty} e^{-\phi^2} b_{2k} d\phi \right) e^{-kR}. \tag{2.30}$$

3. The first terms of the expansion. By calculation we obtain

$$b_0 = 1 \quad b_2 = \frac{R^3 + 6R^2 + 7R + 1}{6R(R+1)^2}\phi^4 - \frac{R^2 + 3R + 1}{9R(R+1)^3}\phi^6. \tag{3.1}$$

Hence from (2.29) we find

$$G_n \sim \pi^{\frac{1}{2}} B \left[1 - \frac{2R^4 + 9R^3 + 16R^2 + 6R + 2}{24Re^R(R+1)^3} \right]. \tag{3.2}$$

Using (2.12) we may write

$$G_n \sim \pi^{\frac{1}{2}} B \left[1 - \frac{2R^4 + 9R^3 + 16R^2 + 6R + 2}{24n(R+1)^3} \right]. \tag{3.3}$$

From (2.4) and (2.15) we have

$$\pi^{\frac{1}{2}} B = n! \exp(e^R - 1)/R^n[2\pi(R+1)\, Re^R]^{\frac{1}{2}}. \tag{3.4}$$

Since $Re^R = n$ we have $e^R = nR^{-1}$, $R^n = n^n e^{-nR}$.

This implies

$$\pi^{\frac{1}{2}} B = \{n! \exp[n(R + R^{-1}) - 1]\}/[2\pi(R+1)]^{\frac{1}{2}} n^{n+\frac{1}{2}}. \tag{3.5}$$

We now use Stirling's expansion for $n!$, namely

$$n! \sim (2\pi)^{\frac{1}{2}} e^{-n} n^{n+\frac{1}{2}} \cdot \left(1 + \frac{1}{12n}\right), \tag{3.6}$$

to obtain from (3.5),

$$\pi^{\frac{1}{2}} B \sim (\exp[n(R + R^{-1} - 1) - 1])(1 + 1/12n)/(R+1)^{\frac{1}{2}}. \tag{3.7}$$

Hence from (3.3) and (3.7) we obtain

$$G_n \sim (R+1)^{-\frac{1}{2}} \exp[n(R + R^{-1} - 1) - 1] \left(1 - \frac{R^2(2R^2 + 7R + 10)}{24n(R+1)^3}\right). \tag{3.8}$$

It would be natural to express R asymptotically in terms of n by means of (2.12) and thus to obtain an asymptotic expression for G_n entirely in terms of n. However, this procedure is not satisfactory as we shall now show. We may rewrite (2.12) in the form

$$R = \log n - \log R. \tag{3.9}$$

Starting with the approximation $R = \log n$ and iterating we find

$$R = \log n - \log(\log n) + \frac{\log(\log n)}{\log n} + \ldots. \tag{3.10}$$

Further terms contain higher powers of $\log n$ in the denominators. However, in (3.8) we have a term of the type nR. It is clear that none of the terms in nR will approach zero as $n \to \infty$. Hence none of these terms can be dropped. For this reason it is better to retain (3.8) as our final result.

This point is overlooked by Epstein **(4)** thus leading to the inaccuracy of (1.3). For $n = 50$, (2.12) gives $R = 2.8608902\ldots$ and (3.8) then yields $G_{50} = 1.85730\ldots \times 10^{47}$. This is an excellent agreement with exact value of G_{50} given in the next section.

4. Table of values of G . Using the recursion formula

$$G_0 = 1, \; G_{n+1} = \sum_{r=0}^{n} \binom{n}{r} G_r \tag{4.1}$$

Bell **(3)** constructed a table of G_n for $n < 20$. By means of an algorithm due to Aitkin **(1)**, H. Finlayson recalculated these values and extended the table up to $n = 25$. At the request of the authors, F. L. Miksa checked these values and calculated the G_n up to $n = 51$. G_{30} was independently calculated by H. W. Becker **(2)**. The Table will be found on page 54.

REFERENCES

1. A. C Aitkin, *A problem on combinations*, Math. Notes Edinburgh, *21* (1933), 18–23.
2. H. W Becker, *Solution to problem*, Amer. Math. Monthly, *48* (1941), 701–702.
3. E. T. Bell, *Exponential polynomials*, Ann. Math., *35* (1934), 258–277.
4. L. F. Epstein, *A function related to the series for e^{e^x}*, J. Math. & Phys., *18* (1939), 166–182.
5. H. Finlayson, *Numbers generated by e^{e^x-1}*. Master's thesis, University of Alberta (1954), unpublished.
6. K. Knopp, *Theory and applications of infinite series* (Blackie and Sons, 1928), p. 563.
7. L. Moser and M. Wyman, *On solutions of $x^d = 1$ in symmetric groups*. Can. J. Math., *7* (1955), 159–168.

TABLE OF G_n $(0 \leqslant n \leqslant 51)$

n	
0	1
1	1
2	2
3	5
4	15
5	52
6	203
7	877
8	4140
9	21147
10	1 15975
11	6 78570
12	42 13597
13	276 44437
14	1908 99322
15	13829 58545
16	1 04801 42147
17	8 28648 69804
18	68 20768 06159
19	583 27422 05057
20	5172 41582 35372
21	47486 98161 56751
22	4 50671 57384 47323
23	44 15200 58550 84346
24	445 95886 92948 05289
25	4638 59033 22299 99353
26	49631 24652 36187 56274
27	5 45717 04793 60599 89389
28	61 60539 40459 99346 52455
29	713 39801 93886 02751 91172
30	8467 49014 51180 93324 50147
31	1 02933 58946 22637 64850 95653
32	12 80646 70049 90871 38189 25644
33	162 95958 92846 00760 67647 28147
34	2119 50393 88640 36046 23886 56799
35	28160 02030 19560 26656 33404 26570
36	3 81971 47298 94818 33997 55256 81317
37	52 86836 62085 50447 90194 55756 24941
38	746 28989 20956 25330 52309 95406 39146
39	10738 82333 07746 92832 76885 79864 25209
40	1 57450 58839 12049 31289 32434 47025 31067
41	23 51152 50774 06176 28200 69407 72437 88988
42	357 42549 19887 26172 91353 50865 66266 42567
43	5529 50118 79716 54843 21714 69328 07377 67385
44	87019 63427 38705 50890 23600 53185 57971 48876
45	13 92585 05266 26366 96023 47053 99365 40796 93415
46	226 54182 19334 49400 29284 84444 70539 22761 58355
47	3745 00595 02461 51119 65053 42096 43151 01201 74682
48	62891 97963 03118 41542 02104 54071 84953 77460 15761
49	10 72613 71545 73358 40034 22155 18590 00263 39172 47281
50	185 72426 87710 78270 43825 77671 81908 91749 92218 52770
51	3263 98387 00041 11524 85695 18301 91582 52441 92558 19477

TRANSACTIONS OF THE ROYAL SOCIETY OF CANADA
VOLUME XLIX : SERIES III : JUNE, 1955
SECTION THREE

Infinite Stochastic Matrices[1]

B. A. RATTRAY and J. E. L. PECK

Presented by R. L. JEFFERY, F.R.S.C.

An $n \times n$ matrix (x_{ij}) is called *doubly stochastic* if for all i and j,

$$x_{ij} \geqslant 0, \sum_i x_{ij} = 1, \sum_j x_{ij} = 1.$$

A (doubly) stochastic matrix in which there is a 1 in each row and column (the other entries therefore 0) is called a *permutation matrix*. A convex mean of permutation matrices is obviously stochastic and many authors (see for example **1**, **2**) have established that these are all of them, according to the following theorem.

The convex set generated by the $n \times n$ permutation matrices is the set of stochastic matrices.

It is of interest to consider the case where the number of rows and columns is (countably) infinite. If an infinite stochastic matrix is such that every non-zero entry is a multiple of a fixed positive number, then a simple application of the solution to the Marriage Problem **(3)** shows that it is a convex mean of permutations **(4)**. However it is easy to find an infinite stochastic matrix which is not a convex mean of permutations, though it is known **(4)** that they are all uniform limits of convex means. As a matter of fact, if we say that a matrix is *sub-stochastic* when it has only non-negative entries and the sum of each row and column is at most 1, then the convex closure of the permutations in the uniform topology is the set of sub-stochastic matrices, and this statement is also true in the topology of point-wise convergence.

Isbell **(4)** has shown that if the topology is strengthened sufficiently, the convex closure of the permutations is a certain subset of the set of stochastic matrices. The topology which he uses is that induced by the norm

$$||x|| = \max(\sup_i \sum_j |x_{ij}|, \sup_j \sum_i |x_{ij}|).$$

Our purpose is to give a topology such that, by a slight modification of Isbell's arguments, we can show that:

The convex closure of the permutations is exactly the set of stochastic matrices.

[1]The work leading to this note was done at the Summer Research Institute of the Canadian Mathematical Congress. The authors wish to thank Professor A. L. Dulmage for arousing their interest in this problem.

Let X be the linear space of all real infinite matrices whose rows and columns are absolutely summable. We define a topology in X by saying that $x \in V_{N,\epsilon}$ if

$$\textstyle\sum_j |x_{ij}| < \epsilon, \ i \leqslant N, \qquad \sum_i |x_{ij}| < \epsilon, \ j \leqslant N.$$

The $V_{N,\epsilon}$ are taken as neighbourhoods of 0. Translations of these neighbourhoods provide X with the topology which we require.

Let x be any stochastic matrix and let $\epsilon > 0$ and N be given. We shall show that there is a stochastic matrix z such that $x - z \in V_{N,\epsilon}$ while every non-zero entry of z is a multiple of a positive number whence z is a convex mean of permutations (see remark above).

To do this choose n so that in each of the first N rows and columns of x, the sum of some n entries is greater than $1 - \epsilon/4$. Choose m so that $n/m < \epsilon/4$. For each x_{ij}, choose an integer p_{ij} so that when $x_{ij} > 0$,

$$p_{ij}/m < x_{ij} \leqslant (p_{ij} + 1)/m$$

and when $x_{ij} = 0$, $p_{ij} = 0$. Put $y_{ij} = p_{ij}/m$. The matrix y is thus substochastic and every row and column of y has a sum of the form $q/m < 1$. Since y has only finitely many non-zero entries in each row and column, we may increase some of the y_{ij} to z_{ij} by an amount $1/m$ so that z is stochastic, by addition of $1/m$ in as many places as is necessary. We adjust the first row then the first column, then the second row and so on; any row can be adjusted because there remains always an infinite number of columns whose sum is not greater than $(m - 1)/m$.

To show that $x - z \in V_{N,\epsilon}$ we consider any one of the first N rows, say. In such a row we may find an index set J of n elements for which

$$\sum_{j \in J} x_{ij} > 1 - \epsilon/4, \quad \sum_{j \notin J} x_{ij} < \epsilon/4.$$

This means that

$$1 > \sum_j y_{ij} \geqslant \sum_{j \in J} y_{ij} \geqslant \sum_{j \in J} \left(x_{ij} - \frac{1}{m}\right) > 1 - \frac{\epsilon}{4} - \frac{n}{m} > 1 - \frac{\epsilon}{2}$$

and

$$\sum_{j \notin J} y_{ij} < \tfrac{1}{2}\epsilon.$$

Thus because $y_{ij} \leqslant x_{ij}$ we have

$$\sum_j |x_{ij} - y_{ij}| \leqslant \sum_{j \in J} (x_{ij} - y_{ij}) + \sum_{j \notin J} x_{ij} < \frac{n}{m} + \frac{\epsilon}{4} < \frac{\epsilon}{2}.$$

Also since $y_{ij} \leqslant z_{ij}$ we have

$$\textstyle\sum_j |z_{ij} - y_{ij}| = \sum_j z_{ij} - \sum_j y_{ij} < 1 - (1 - \frac{1}{2}\epsilon) = \frac{1}{2}\epsilon.$$

This means that

$$\textstyle\sum_j |x_{ij} - z_{ij}| \leqslant \sum_j |x_{ij} - y_{ij}| + \sum_j |y_{ij} - z_{ij}| < \epsilon.$$

But since this can be done for any of the first N rows and columns, we have that $x - z \in V_{N,\epsilon}$. This establishes that every stochastic matrix is a limit of convex means of permutations. It is not difficult to verify that every limit of convex means of permutations is stochastic and this completes the proof.

It is interesting to notice that if our topology is strenghened slightly by requiring uniform convergence in addition, then the same conclusion holds.

References

1. G. Birkhoff, *Tres observaciones sobre el algebra lineal*, Rev., Mat. y Fisica Teorica, Tucuman Ser A *5* (1946), 147–150.
2. A. J. Hoffman and H. W. Wielandt, *The variation of the spectrum of a normal matrix*, Duke Math. J., *20* (1953), 37–39.
3. P. R. Halmos and H. E. Vaughan, *The marriage problem*, Amer. J. Math., *72* (1950), 214–215.
4. J. R. Isbell, *Birkhoff's Problem 111*, Proc. Amer. Math. Soc., *6* (1955), 217–218.

TRANSACTIONS OF THE ROYAL SOCIETY OF CANADA
VOLUME XLIX : SERIES III : JUNE, 1955
SECTION THREE

A Generalization of Some Theorems of Hardy

P. G. ROONEY

Presented by MAX WYMAN, F.R.S.C.

1. Introduction. G. H. Hardy **(1, 2)** has shown that if $\phi \in L_p(0, \infty)$, and

$$\text{I} \qquad \xi(x) = \frac{1}{x}\int_0^x \phi(t)\,dt,$$

$$\text{II} \qquad \eta(x) = \int_x^\infty \frac{\phi(t)}{t}\,dt,$$

then $\xi \in L_p(0, \infty)$ for $1 < p \leqslant \infty$, and $\eta \in L_p(0, \infty)$ for $1 \leqslant p < \infty$.

The spaces $L_p(0, \infty)$, $1 \leqslant p < \infty$, have been generalized by Lorentz **(3)** to spaces $\Lambda(\alpha, p)$ in the following manner. Let $\alpha > 0$, $p \geqslant 1$; let ϕ be measurable on $(0, \infty)$, and let $|\phi|$ possess there an equi-measurable decreasing rearrangement ϕ^*. Define

$$\text{III} \qquad ||\phi(\cdot)||_{\Lambda(\alpha,p)} = \left\{\alpha\int_0^\infty x^{\alpha-1}(\phi^*(x))^p dx\right\}^{1/p}.$$

Then $\Lambda(\alpha, p)$ consists of those ϕ for which this norm is finite. It is a Banach space if $0 < \alpha \leqslant 1$.

The role of $L_\infty(0, \infty)$ as conjugate space of $L_1(0, \infty)$ is played in this theory, at least for $0 < \alpha < 1$, by the space $M(\alpha, 1)$ which is a special case of the spaces $M(\alpha, p)$ defined for $0 \leqslant \alpha \leqslant 1$, to consist of those ϕ measurable on $(0, \infty)$ for which $|\phi|$ possess there an equi-measurable rearrangement ϕ^* and for which the norm

$$\text{IV} \qquad ||\phi(\cdot)||_{M(\alpha,p)} = \sup_e (m(e))^{-\alpha}\left\{\int_e |\phi(x)|^p dx\right\}^{1/p}, \quad e \subset (0, \infty),\ m(e) < \infty,$$

is finite. $M(\alpha, p)$ is a Banach space for each α, $0 \leqslant \alpha \leqslant 1$.

Our object here is to generalize Hardy's theorems to the spaces $\Lambda(\alpha, p)$ and $M(\alpha, p)$. This generalization forms the contents of section three. The generalizations are facilitated by a theorem on the structure of $M(\alpha, p)$ which is proved in section two.

Hardy's results have been applied to show that if $\phi \in L_p(0, \infty)$, $1 < p < \infty$, and

$$\text{V} \qquad f(s) = \int_0^\infty e^{-st}\phi(t)\,dt = \mathscr{L}(\phi(\cdot); s),$$

then $s^{1-(2/p)} f(s) \in L_p(0, \infty)$. A proof of this is given in **(5**; p. 397**)**. We shall give similar applications of our theorems in section four.

2. A structure theorem. It is clear that $\Lambda(1, p) = \mathrm{M}(0, p) = L_p(0, \infty)$. Further information is provided by the following theorem.

THEOREM 1. *If $p\alpha > 1$, $\mathrm{M}(\alpha, p)$ is void. Further if $p\alpha = 1$, $\mathrm{M}(\alpha, p) = L_\infty(0, \infty)$.*

Proof. From the definition of the norm, it is clear that if $\phi \in \mathrm{M}(\alpha, p)$, $e \subset (0, \infty)$, $m(e) < \infty$,

$$\left\{ \int_e |\phi(x)|^p dx \right\}^{1/p} \leqslant (m(e))^\alpha ||\phi(\cdot)||_{\mathrm{M}(\alpha,p)}.$$

Hence in particular, if $0 < |h| \leqslant x$,

$$\left\{ \frac{1}{h} \int_x^{x+h} |\phi(t)|^p dt \right\}^{1/p} \leqslant h^{\alpha - 1/p} ||\phi(\cdot)||_{\mathrm{M}(\alpha,p)}.$$

But for almost all $x \geqslant 0$,

$$\lim_{h \to 0} \frac{1}{h} \int_x^{x+h} |\phi(t)|^p dt = |\phi(x)|^p.$$

Hence if $p\alpha > 1$, then almost everywhere we have

$$|\phi(x)| \leqslant \overline{\lim_{h \to 0}}\, h^{\alpha - 1/p} ||\phi(\cdot)||_{\mathrm{M}(\alpha,p)} = 0,$$

and $\mathrm{M}(\alpha, p)$ is void. If $p\alpha = 1$, then for almost all $x \geqslant 0$,

$$|\phi(x)| \leqslant ||\phi(\cdot)||_{\mathrm{M}(\alpha,p)},$$

and $\phi \in L_\infty(0, \infty)$, and

$$||\phi(\cdot)||_{L_\infty} \leqslant ||\phi(\cdot)||_{\mathrm{M}(\alpha,p)}.$$

Conversely, if $p\alpha = 1$, and $\phi \in L_\infty(0, \infty)$, then

$$||\phi(\cdot)||_{\mathrm{M}(\alpha,p)} = \sup_e (m(e))^{-\alpha} \left\{ \int_e |\phi(x)|^p dx \right\}^{1/p}$$

$$\leqslant ||\phi(\cdot)||_{L_\infty} \sup_e (m(e))^{-\alpha} \left\{ \int_e dx \right\}^{1/p} = ||\phi(\cdot)||_{L_\infty},$$

and $\phi \in \mathrm{M}(\alpha, p)$. Further it is clear that $||\phi(\cdot)||_{L_\infty} = ||\phi(\cdot)||_{\mathrm{M}(\alpha,p)}$.

3. Main theorems. Theorems 2 and 3 deal with ξ, and 4 and 5 with η.

THEOREM 2. *If $\phi \in \Lambda(\alpha, p)$, $1 \leqslant p < \infty$, $0 < \alpha < p$, then $\xi \in \Lambda(\alpha, p)$*

Proof. Clearly if $x > 0$,

$$|\xi(x)| \leqslant \frac{1}{x} \int_0^x |\phi(t)|\, dt \leqslant \frac{1}{x} \int_0^x \phi^*(t)\, dt = \psi(x).$$

But since ϕ^* is non-increasing, ψ is non-increasing, and thus $\xi^*(x) \leqslant \psi(x)$, and it suffices to show that $\psi \in \Lambda(\alpha, p)$.

Now if $p = 1$, this is easy. For then $0 < \alpha < 1$, and hence by Fubini's theorem,

$$||\psi(\cdot)||_{\Lambda(\alpha,1)} = \alpha \int_0^\infty x^{\alpha-1}\psi(x)\,dx = \alpha \int_0^\infty x^{\alpha-2}dx \int_0^x \phi^*(t)\,dt$$

$$= \alpha \int_0^\infty \phi^*(t)\,dt \int_t^\infty x^{\alpha-2}dx = \frac{\alpha}{1-\alpha}\int_0^\infty t^{\alpha-1}\phi^*(t)\,dt = \frac{1}{1-\alpha}||\phi(\cdot)||_{\Lambda(\alpha,1)} < \infty.$$

Suppose then that $1 < p < \infty$, and let $\psi_1(x) = x\psi(x)$. Then by Hölder's inequality, for $x > 0$,

$$\psi_1(x) = \int_0^x \phi^*(t)\,dt \leqslant \left\{\int_0^x t^{\alpha-1}(\phi^*(t))^p dt\right\}^{1/p}\left\{\int_0^x t^{(1-\alpha)/(p-1)}dt\right\}^{(p-1)/p}$$

$$= \left(\frac{p-1}{p-\alpha}\right)^{(p-1)/p} x^{(p-\alpha)/p}\left\{\int_0^x t^{\alpha-1}(\phi^*(t))^p dt\right\}^{1/p}.$$

Thus

$$x^{\alpha-p}(\psi_1(x))^p = \left(\frac{p-1}{p-\alpha}\right)^{p-1}\int_0^x t^{\alpha-1}(\phi^*(t))^p dt,$$

so that

$$x^{\alpha-p}(\psi_1(x))^p = \begin{cases} o(1) & \text{as } x \to 0+ \\ O(1) & \text{as } x \to \infty. \end{cases}$$

But clearly, if $0 < a < b < \infty$,

$$\alpha \int_a^b x^{\alpha-1}(\psi(x))^p dx < \infty.$$

Further, integrating by parts and applying Hölder's inequality,

$$\alpha \int_a^b x^{\alpha-1}(\psi(x))^p dx = \alpha \int_a^b x^{\alpha-p-1}(\psi_1(x))^p dx$$

$$= \frac{\alpha}{p-\alpha}\left\{a^{\alpha-p}(\psi_1(a))^p - b^{\alpha-p}(\psi_1(b))^p\right\} + \frac{\alpha p}{p-\alpha}\int_a^b x^{\alpha-p}(\psi_1(x))^{p-1}\phi^*(x)\,dx$$

$$\leqslant \frac{\alpha}{p-\alpha}a^{\alpha-p}(\psi_1(a))^p + \frac{\alpha p}{p-\alpha}\int_a^b x^{\alpha-1}(\psi(x))^{p-1}\phi^*(x)\,dx$$

$$\leqslant o(1) + \frac{p}{p-\alpha}\left\{\alpha\int_a^b x^{\alpha-1}(\phi^*(x))^p dx\right\}^{1/p}\left\{\alpha\int_a^b x^{\alpha-1}(\psi(x))^p dx\right\}^{(p-1)/p}.$$

Hence, dividing by the last factor of the last term of this expression, and letting $a \to 0+$, $b \to \infty$, we obtain

$$\left\{\alpha\int_0^\infty x^{\alpha-1}(\psi(x))^p dx\right\}^{1/p} \leqslant \frac{p}{p-\alpha}\left\{\alpha\int_0^\infty x^{\alpha-1}(\phi^*(x))^p dx\right\}^{1/p};$$

that is,

$$||\psi(\cdot)||_{\Lambda(\alpha,p)} \leqslant \frac{p}{p-\alpha}||\phi(\cdot)||_{\Lambda(\alpha,p)}$$

and $\psi \in \Lambda(\alpha, p)$.

THEOREM 3. *If* $\phi \in \mathrm{M}(\alpha, p)$, $0 \leqslant \alpha \leqslant 1$, $p \geqslant 1$, *then* $\xi \in \mathrm{M}(\alpha, p)$, *except in the case* $\alpha = 0$, $p = 1$.

Proof. If $p\alpha > 1$, $\mathrm{M}(\alpha, p)$ is void and the theorem is trivially true. If $p\alpha = 1$, then $\mathrm{M}(\alpha, p) = L_\infty(0, \infty)$, and the theorem is clearly true since

$$|\xi(x)| \leqslant \frac{1}{x}\int_0^x |\phi(t)|\, dt \leqslant ||\phi(\cdot)||_{L_\infty}.$$

If $\alpha = 0$, $p > 1$, then $\mathrm{M}(\alpha, p) = L_p(0, \infty)$, and the theorem follows from Hardy (1). Hence we may assume $0 < \alpha < 1, p\alpha < 1$.

Clearly, for $p = 1$,

$$\xi^*(x) \leqslant \frac{1}{x}\int_0^x \phi^*(t)\, dt \leqslant x^{\alpha-1}||\phi(\cdot)||_{\mathrm{M}(\alpha,1)}.$$

If $p > 1$, then by Hölder's inequality

$$\xi^*(x) \leqslant \frac{1}{x}\int_0^x \phi^*(t)\, dt \leqslant \frac{1}{x}\left\{\int_0^x (\phi^*(t))^p dt\right\}^{1/p}\left\{\int_0^x dt\right\}^{(p-1)/p}$$
$$= x^{-1/p}\left\{\int_0^x (\phi^*(t))^p dt\right\}^{1/p} \leqslant x^{\alpha-1/p}||\phi(\cdot)||_{\mathrm{M}(\alpha,p)}.$$

Thus for $1 \leqslant p < \infty$,

$$\xi^*(x) \leqslant x^{\alpha-1/p}||\phi(\cdot)||_{\mathrm{M}(\alpha,p)},$$

and hence if $e \subset (0, \infty)$, $m(e) = \delta < \infty$,

$$\delta^{-\alpha}\left\{\int_e |\xi(x)|^p dx\right\}^{1/p} \leqslant \delta^{-\alpha}\left\{\int_0^\delta (\xi^*(x))^p dx\right\}^{1/p}$$
$$\leqslant ||\phi(\cdot)||_{\mathrm{M}(\alpha,p)}\delta^{-\alpha}\left\{\int_0^\delta x^{p\alpha-1} dx\right\}^{1/p} = (p\alpha)^{-1/p}||\phi(\cdot)||_{\mathrm{M}(\alpha,p)}.$$

so that

$$||\xi(\cdot)||_{\mathrm{M}(\alpha,p)} \leqslant (p\alpha)^{-1/p}||\phi(\cdot)||_{\mathrm{M}(\alpha,p)} < \infty.$$

THEOREM 4. *If* $\phi \in \Lambda(\alpha, p), 0 < \alpha \leqslant 1, p \geqslant 1$, *then* $\eta \in \Lambda(\alpha, p)$.

Proof. Clearly if $x > 0$,

$$|\eta(x)| \leqslant \int_x^\infty \frac{|\phi(t)|}{t}\, dt = \lambda(x).$$

But $\lambda(x)$ is plainly non-increasing, and thus $\eta^*(x) \leqslant \lambda(x)$, and it suffices to show $\lambda \in \Lambda(\alpha, p)$.

Now for $p = 1$ this is easy. For by Fubini's theorem,

$$||\lambda(\cdot)||_{\Lambda(\alpha,1)} = \alpha\int_0^\infty x^{\alpha-1}\lambda(x)\, dx = \alpha\int_0^\infty x^{\alpha-1}dx\int_x^\infty \frac{|\phi(t)|}{t}\, dt$$
$$= \alpha\int_0^\infty \frac{|\phi(t)|}{t}\, dt\int_0^t x^{\alpha-1}dx$$
$$= \int_0^\infty t^{\alpha-1}|\phi(t)|\, dt \leqslant \int_0^\infty t^{\alpha-1}\phi^*(t)\, dt = \frac{1}{\alpha}\, ||\phi(\cdot)||_{\Lambda(\alpha,1)} < \infty.$$

Suppose then $1 < p < \infty$. Then by Hölder's inequality, if $x > 0$,

$$\lambda(x) = \int_x^\infty \frac{|\phi(t)|}{t}\,dt \leqslant \left\{\int_x^\infty t^{\alpha-1}|\phi(t)|^p dt\right\}^{1/p}\left\{\int_x^\infty t^{-(p+\alpha-1)/(p-1)}dt\right\}^{(p-1)/p}$$
$$= \left(\frac{p-1}{\alpha}\right)^{(p-1)/p} x^{-\alpha/p}\left\{\int_x^\infty t^{\alpha-1}|\phi(t)|^p dt\right\}^{1/p}.$$

Hence

$$x^\alpha(\lambda(x))^p \leqslant \left(\frac{p-1}{\alpha}\right)^{p-1}\int_x^\infty t^{\alpha-1}|\phi(t)|^p dt,$$

so that

$$x^\alpha(\lambda(x))^p = o(1) \text{ as } x \to \infty$$
$$= O(1) \text{ as } x \to 0+.$$

Now clearly if $0 < a < b < \infty$,

$$\alpha\int_a^b x^{\alpha-1}(\lambda(x))^p dx < \infty.$$

Further, integrating by parts and using Hölder's inequality,

$$\alpha\int_a^b x^{\alpha-1}(\lambda(x))^p dx = \{b^\alpha(\lambda(b))^p - a^\alpha(\lambda(a))^p\} + p\int_a^b x^{\alpha-1}(\lambda(x))^{p-1}|\phi(x)|\,dx$$
$$\leqslant b^\alpha(\lambda(b))^p + \frac{p}{\alpha}\left\{\alpha\int_a^b x^{\alpha-1}|\phi(x)|^p dx\right\}^{1/p}\left\{\alpha\int_a^b x^{\alpha-1}(\lambda(x))^p dx\right\}^{(p-1)/p}.$$

Hence, dividing by the last factor in the last term of this expression, and letting $a \to 0+$, $b \to \infty$, we obtain

$$\left\{\alpha\int_0^\infty x^{\alpha-1}(\lambda(x))^p dx\right\}^{1/p} \leqslant \frac{p}{\alpha}\left\{\alpha\int_0^\infty x^{\alpha-1}|\phi(x)|^p dx\right\}^{1/p}$$
$$\leqslant \frac{p}{\alpha}\left\{\int_0^\infty x^{\alpha-1}(\phi^*(x))^p dx\right\}^{1/p};$$

that is

$$||\lambda(\cdot)||_{\Lambda(\alpha,p)} \leqslant \frac{p}{\alpha}||\phi(\cdot)||_{\Lambda(\alpha,p)},$$

and $\lambda \in \Lambda(\alpha, p)$.

THEOREM 5. *If* $\phi \in \mathrm{M}(\alpha, p)$, $0 \leqslant \alpha \leqslant 1$, $p \geqslant 1$, $p\alpha \neq 1$, *then* $\eta \in \mathrm{M}(\alpha, p)$.

Proof. Since $\mathrm{M}(\alpha, p)$ is void for $p\alpha > 1$, we may assume $p\alpha < 1$. Further since $\mathrm{M}(0, p) = L_p(0, \infty)$, and the theorem is true in this case by Hardy **(2)**, we may assume $0 < \alpha < 1$.

We show first that if $p\alpha \leqslant 1$,

$$\int_0^t |\phi(u)|\,du \leqslant t^{\alpha+(p-1)/p}||\phi(\cdot)||_{\mathrm{M}(\alpha,p)}.$$

For $p = 1$, this follows immediately from the definition of $||\phi(\cdot)||_{M(\alpha,p)}$. For $p > 1$, we obtain from Hölder's inequality,

$$\int_0^t |\phi(u)|\, du \leqslant \left\{\int_0^t |\phi(u)|^p du\right\}^{1/p} \left\{\int_0^t du\right\}^{(p-1)/p}$$

$$= t^{(p-1)/p}\left\{\int_0^t |\phi(u)|^p du\right\}^{1/p} \leqslant t^{\alpha+(p-1)/p}||\phi(\cdot)||_{M(\alpha,p)}.$$

Further, as in the proof of the previous theorem,

$$\eta^*(x) \leqslant \int_x^\infty \frac{|\phi(t)|}{t}\, dt.$$

Hence, integrating by parts,

$$\eta^*(x) \leqslant \frac{1}{t}\int_0^t |\phi(u)|\, du \Bigg|_x^\infty + \int_x^\infty \frac{dt}{t^2}\int_0^t |\phi(u)|\, du.$$

Now since $p\alpha < 1$, the first of these terms is less than

$$\overline{\lim_{R\to\infty}} \frac{1}{R}\int_0^R |\phi(u)|\, du - \frac{1}{x}\int_0^x |\phi(x)|\, dx \leqslant \overline{\lim_{R\to\infty}} R^{\alpha-1/p}||\phi(\cdot)||_{M(\alpha,p)} = 0,$$

and the second is less than

$$||\phi(\cdot)||_{M(\alpha,p)}\int_x^\infty t^{\alpha-1-1/p}dt = \frac{p}{1-p\alpha}x^{\alpha-1/p}||\phi(\cdot)||_{M(\alpha,p)}.$$

Hence

$$\eta^*(x) \leqslant \frac{p}{1-p\alpha}x^{\alpha-1/p}||\phi(\cdot)||_{M(\alpha,p)},$$

and thus if $e \subset (0, \infty)$, $m(e) = \delta < \infty$,

$$\delta^{-\alpha}\left\{\int_e |\eta(x)|^p dx\right\}^{1/p} \leqslant \delta^{-\alpha}\left\{\int_0^\delta (\eta^*(x))^p dx\right\}^{1/p}$$

$$\leqslant \frac{p}{1-p\alpha}\, ||\phi(\cdot)||_{M(\alpha,p)}\delta^{-\alpha}\left\{\int_0^\delta x^{p\alpha-1}dx\right\}^{1/p}$$

$$= \frac{p}{1-p\alpha}\,(p\alpha)^{-1/p}||\phi(\cdot)||_{M(\alpha,p)},$$

so that

$$||\eta(\cdot)||_{M(\alpha,p)} \leqslant \frac{p}{1-p\alpha}\,(p\alpha)^{-1/p}||\phi(\cdot)||_{M(\alpha,p)},$$

and $\eta \in M(\alpha, p)$.

3. An application to Laplace integrals. The result mentioned in the introduction, that if $\phi \in L_p(0, \infty)$, $1 < p \leqslant \infty$, and $f(s) = \mathcal{L}(\phi(\cdot); s)$, then $s^{-1(2/p)}f(s) \in L_p(0, \infty)$, is equivalent to $s^{-1}f(s^{-1}) \in L_p(0, \infty)$. In this form we shall generalize the result to $\Lambda(\alpha, p)$ and $M(\alpha, p)$.

The results given in the corollaries to the two following theorems are also to be found, for $p = 1, 0 < \alpha < 1$, in **(4)**.

THEOREM 6. *If* $\phi \in \Lambda(\alpha, p)$, $1 \leqslant p < \infty$, $0 < \alpha < 1$, *and* $f(s) = \mathscr{L}(\phi(\cdot); s)$, *then* $s^{-1} f(s^{-1}) \in \Lambda(\alpha, p)$.

Proof. Clearly, if $s > 0$,

$$s^{-1}f(s^{-1}) \leqslant \frac{1}{s}\int_0^s e^{-x/s}|\phi(x)|\,dx + \frac{1}{s}\int_s^\infty e^{-x/s}|\phi(x)|\,dx$$
$$\leqslant \frac{1}{s}\int_0^s |\phi(x)|\,dx + \int_s^\infty \frac{|\phi(x)|}{x}\,dx = \xi_1(s) + \eta_1(x).$$

But by Theorems 2 and 4, ξ_1 and $\eta_1 \in \Lambda(\alpha, p)$. Hence

$$||s^{-1}f(s^{-1})||_{\Lambda(\alpha,p)} \leqslant ||\xi_1(\cdot) + \eta_1(\cdot)||_{\Lambda(\alpha,p)}$$
$$\leqslant ||\xi_1(\cdot)||_{\Lambda(\alpha,p)} + ||\eta_1(\cdot)||_{\Lambda(\alpha,p)} < \infty,$$

and $s^{-1} f(s^{-1}) \in \Lambda(\alpha, p)$.

COROLLARY. *If* $\phi \in \Lambda(\alpha, p)$, $1 \leqslant p < \infty$, $0 < \alpha < 1$, *then*

$$\int_0^\infty s^{\alpha-p-1}|f(s^{-1})|^p ds < \infty.$$

Proof.
$$\int_0^\infty s^{\alpha-p-1}|f(s^{-1})|^p ds = \int_0^\infty s^{\alpha-1}|s^{-1}f(s^{-1})|^p ds$$
$$\leqslant \int_0^\infty s^{\alpha-1}(s^{-1}f(s^{-1}))^{*p} ds = ||s^{-1}f(s^{-1})||_{\Lambda(\alpha,p)} < \infty.$$

THEOREM 7. *If* $\phi \in \mathrm{M}(\alpha, p)$, $0 \leqslant \alpha \leqslant 1$, $1 \leqslant p < \infty$, *then, except in the case* $\alpha = 0, p = 1$, $s^{-1} f(s^{-1}) \in \mathrm{M}(\alpha, p)$.

Proof. A proof along the lines of the proof of Theorem 6 is easy to construct. However a direct attack yields more information.

We may confine ourselves to the case $0 < \alpha < 1$, $p\alpha < 1$. For if $p\alpha > 1$, $\mathrm{M}(\alpha, p)$ is void, and if $p\alpha = 1$, $\mathrm{M}(\alpha, p) = L_\infty(0, \infty)$, in which case the theorem is trivial. If $\alpha = 0$, $\mathrm{M}(\alpha, p) = L_p(0, \infty)$ and, as mentioned in the introduction, the theorem for this case is known.

We note first that, as in the proof of Theorem 5,

$$\int_0^t |\phi(u)|\,du \leqslant t^{\alpha+(p-1)/p}||\phi(\cdot)||_{\mathrm{M}(\alpha,p)}.$$

Then integrating by parts we have for $s > 0$,

$$|f(s)| \leqslant \int_0^\infty e^{-st}|\phi(t)|\,dt = e^{-st}\int_0^t |\phi(u)|\,du\,\Big|_0^\infty$$
$$+ s\int_0^\infty e^{-st}dt\int_0^t |\phi(u)|\,du.$$

But the first of these terms is less than

$$\overline{\lim_{R\to\infty}}\, e^{-sR}\int_0^R |\phi(t)|\, dt \leqslant \overline{\lim_{R\to\infty}}\, e^{-sR} R^{\alpha+(p-1)/p}||\phi(\cdot)||_{\mathbf{M}(\alpha,p)} = 0,$$

and the second is less than

$$s||\phi(\cdot)||_{\mathbf{M}(\alpha,p)}\int_0^\infty e^{-st}t^{\alpha+(p-1)/p}dt = s^{-(\alpha+(p-1)/p)}\Gamma(\alpha+2-1/p)||\phi(\cdot)||_{\mathbf{M}(\alpha,p)}.$$

Hence,

$$s^{-1}|f(s^{-1})| \leqslant s^{\alpha-1/p}\,\Gamma(\alpha+2-1/p)||\phi(\cdot)||_{\mathbf{M}(\alpha,p)}.$$

But since $p\alpha < 1$, this last function is decreasing, and thus

$$(s^{-1}f(s^{-1}))^* \leqslant s^{\alpha-1/p}\,\Gamma(\alpha+2-1/p)||\phi(\cdot)||_{\mathbf{M}(\alpha,p)}.$$

Thus if $e \subset (0, \infty)$, $m(e) = \delta < \infty$,

$$\begin{aligned}\delta^{-\alpha}\left\{\int_e |s^{-1}f(s^{-1})|^p ds\right\}^{1/p} &\leqslant \delta^{-\alpha}\left\{\int_0^\delta (s^{-1}f(s^{-1}))^{*p} ds\right\}^{1/p}\\ &\leqslant \Gamma(\alpha+2-1/p)||\phi(\cdot)||_{\mathbf{M}(\alpha,p)}\delta^{-\alpha}\left\{\int_0^\delta s^{p\alpha-1}ds\right\}^{1/p}\\ &= (p\alpha)^{-1/p}\,\Gamma(\alpha+2-1/p)||\phi(\cdot)||_{\mathbf{M}(\alpha,p)} < \infty,\end{aligned}$$

and hence

$$||s^{-1}f(s^{-1})||_{\mathbf{M}(\alpha,p)} < \infty.$$

COROLLARY. *If $\phi \in \mathrm{M}(\alpha, p)$, $0 \leqslant \alpha \leqslant 1$, $1 \leqslant p < \infty$, and $f(s) = \mathcal{L}(\phi(\cdot); s)$, then $s^{-(\alpha+1-1/p)}f(s^{-1})$ is bounded.*

Proof. This was proved in the course of the previous theorem.

REFERENCES

1. G. H. Hardy, *An inequality between integrals*, Messenger of Math., *54* (1925), 150–156.
2. G. H. Hardy, *Further inequalities between integrals*, Messenger of Math., *57* (1927), 12–16.
3. G. G. Lorentz, *Some new functional spaces*, Ann. Math., *51* (1950), 37–55.
4. P. G. Rooney, *An application of some spaces of Lorentz*, Can. J. Math., *7* (1955), 314–321.
5. E. C. Titchmarsh, *Theory of functions* (2nd ed., Oxford, 1939).

PRINTED IN CANADA
AT THE UNIVERSITY OF TORONTO PRESS